

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 31.9642 Seconds
(without alignments)
1763.712 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKNNLLTKKPIANKSN.....PAGLGALLGRRKRKNKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/ias/5A COMB.pep.*
2: /cgn2_6/ptodata/2/ias/5B COMB.pep.*
3: /cgn2_6/ptodata/2/ias/6A COMB.pep.*
4: /cgn2_6/ptodata/2/ias/6B COMB.pep.*
5: /cgn2_6/ptodata/2/ias/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/ias/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4490	79.5	930	4	US-09-134-001C-5314
2	4485	79.4	930	4	US-09-386-962C-10
3	2698.5	47.8	1166	4	US-09-200-6508-7
4	2303.5	40.8	1742	4	US-09-386-962C-4
5	1982	35.1	930	4	US-09-200-6508-3
6	1880	33.3	1315	4	US-09-200-6508-5
7	1629.5	28.9	936	4	US-08-956-171B-5249
8	1628	28.8	933	3	US-08-933-728-2
9	1628	28.8	933	3	US-09-421-868-2
10	1613.5	28.6	918	4	US-09-200-6508-1
11	671.5	11.9	487	4	US-09-386-962C-14
12	635.5	11.3	2137	4	US-09-134-001C-4463
13	488	8.6	1027	4	US-08-956-171B-5254
14	482	8.5	466	4	US-09-134-001C-4749
15	408.5	7.2	669	4	US-09-107-532A-6532
16	377.5	6.7	1183	4	US-09-134-001C-3530
17	349	6.2	345	3	US-08-856-253-7
18	331.5	5.9	1161	4	US-09-327-536-2
19	315.5	5.6	1112	2	US-08-714-402-2
20	301.5	5.3	1231	3	US-08-904-263A-4
21	301.5	5.3	1231	4	US-09-434-123A-4
22	295	5.2	2315	4	US-09-543-681A-5434
23	292.5	5.2	886	4	US-08-956-171B-5235
24	292.5	5.2	2504	4	US-09-328-352-5821
25	286	5.1	1060	4	US-08-911-393-2
26	282	5.0	3052	2	US-08-557-122A-26
27	282	5.0	3052	4	US-09-262-666-26

28	279	4.9	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
29	278	4.9	699	4	US-09-134-001C-4054	Sequence 4054, Ap
30	276.5	4.9	1085	1	US-08-431-080-28	Sequence 28, Appl
31	276.5	4.9	1085	2	US-08-938-534-28	Sequence 28, Appl
32	276.5	4.9	1085	4	US-09-345-294-28	Sequence 28, Appl
33	276	4.9	2391	2	US-08-446-855A-2	Sequence 2, Appl
34	276	4.9	2391	3	US-09-150-741-2	Sequence 2, Appl
35	271	4.8	1115	2	US-08-568-459A-2	Sequence 2, Appl
36	271	4.8	1115	2	US-08-487-826B-2	Sequence 2, Appl
37	271	4.8	1115	4	US-09-210-288-2	Sequence 2, Appl
38	271	4.8	1115	6	5198347-6	Patent No. 5198347
39	270.5	4.8	1833	4	US-08-621-944A-4	Sequence 4, Appl
40	270.5	4.8	1833	4	US-08-945-567D-4	Sequence 4, Appl
41	270.5	4.8	1992	4	US-08-621-944A-3	Sequence 3, Appl
42	270.5	4.8	1992	4	US-08-945-567D-3	Sequence 3, Appl
43	269.5	4.8	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
44	269	4.8	60	4	US-09-386-962C-19	Sequence 19, Appl
45	267	4.7	2123	3	US-08-968-685A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-134-001C-5314

; Sequence 5314, Application US/09134001C

; Patent No. 5380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5314

; LENGTH: 930

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5314

Query Match 79.5%; Score 4490; DB 4; Length 930;

Best Local Similarity 81.2%; Pred. No. 1.1e-241;

Matches 884; Conservative 19; Mismatches 26; Indels 160; Gaps 3;

QY	4	KGNLLTKKPIANKSNKYAIRKFTVTGTAIVIGATLLFGLGHNEAKAEENSVDQVDSN	63
DB	2	KGNLLTKKPIANKSNKYAIRKFTVTGTAIVIGATLLFGLGHNEAKAEENTVDQVDSN	61
QY	64	TDELSDSNDQSDDEKNDVINNQSDINTDDNQIIKKEETNNVDGIEKSEDETESTTN	123
DB	62	MODELSDSNDQSDNEEKNDVINNQSDINTDDNQ-IKKEETNSNDATENRKSQITQSTTN	120
QY	124	VDENEATFLQKTPQDMTHLTERRVKSSSVESSENSISDTAQPSHTTINREESVQTSQDNV	183
DB	121	VDENEATFLQKTPQDMTHLTERRVKSSSVESSENSISDTAQPSHTTINREESVQTSQDNV	180
QY	184	EDSHVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN	243
DB	181	ENSRVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN	240
QY	244	LPINEYENKARPLSTTSQAQPSKRVTVNQAABQGSNVNHLIKVTDQSTIEGYDDSEGI	303
DB	241	LPINEYENKARPLSTTSQAQPSKRVTVNQAABQGSNVNHLIKVTDQSTIEGYDDSEGI	300
QY	304	KAHDAENLIYDVFTEKVDKVGSDTMTVDIDKNTVPSDLTDSFTIPKIKONSSEIATGT	363
DB	301	KAHDAENLIYDVFTEKVDKVGSDTMTVDIDKNTVPSDLTDSFTIPKIKONSSEIATGT	360

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 128.509 Seconds
(without alignments)
2734.476 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKKNLLTKKKPIANKSN.....FAGLGALLGRRKRNKNKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	15	US-10-378-674-8
2	4485	79.4	930	12	US-10-615-383-10
3	4485	79.4	930	16	US-10-689-082-10
4	4485	79.4	930	16	US-10-690-184-10
5	4309	76.3	892	16	US-10-661-809-21
6	4307	76.3	892	12	US-10-282-122A-70481
7	3351.5	59.4	670	12	US-10-282-122A-70444
8	2787	49.4	549	15	US-10-378-674-9
9	2624.5	46.5	560	15	US-10-378-674-2
10	2586	45.8	1141	12	US-10-282-122A-70251
11	2339.5	41.4	1633	12	US-10-282-122A-70437
12	2303.5	40.8	1742	12	US-10-615-383-4
13	2303.5	40.8	1742	16	US-10-689-082-4
14	2303.5	40.8	1742	16	US-10-690-184-4
15	2234	39.6	1385	12	US-10-282-122A-44324

16	2205.5	39.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
17	2093.5	37.1	953	12	US-10-282-122A-44457	Sequence 44457, A
18	2069	36.6	1349	9	US-09-815-242-5898	Sequence 5898, Ap
19	2069	36.6	1349	9	US-09-815-242-13137	Sequence 13137, A
20	2011	35.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
21	2011	35.6	932	9	US-09-815-242-12438	Sequence 12438, A
22	1672	29.6	343	15	US-10-378-674-4	Sequence 4, Appli
23	1634	28.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
24	1634	28.9	1021	9	US-09-815-242-12544	Sequence 12544, A
25	1629.5	28.9	936	8	US-08-781-986A-5249	Sequence 5249, Ap
26	1629.5	28.9	936	12	US-10-329-624-5249	Sequence 5249, Ap
27	1581	28.0	935	12	US-10-282-122A-44326	Sequence 44326, A
28	1562	27.7	316	15	US-10-378-674-6	Sequence 6, Appli
29	1536	27.2	877	12	US-10-282-122A-70428	Sequence 70428, A
30	1441.5	25.5	841	9	US-09-815-242-5779	Sequence 5779, Ap
31	1441.5	25.5	841	9	US-09-815-242-12751	Sequence 12751, A
32	1121.5	19.9	513	12	US-10-282-122A-59154	Sequence 59154, A
33	1023	18.1	265	13	US-10-073-256-78	Sequence 78, Appli
34	865	15.3	1113	9	US-09-815-242-5836	Sequence 5836, Ap
35	836	14.8	897	9	US-09-815-242-12769	Sequence 12769, A
36	782.5	13.9	903	12	US-10-282-122A-60847	Sequence 60847, A
37	708.5	12.5	1253	14	US-10-363-798-2	Sequence 2, Appli
38	671.5	11.9	487	12	US-10-615-383-14	Sequence 14, Appli
39	671.5	11.9	487	16	US-10-689-082-14	Sequence 14, Appli
40	671.5	11.9	487	16	US-10-690-184-14	Sequence 14, Appli
41	624.5	11.1	2271	12	US-10-282-122A-43924	Sequence 43924, A
42	615.5	10.9	2283	14	US-10-172-502-4	Sequence 4, Appli
43	607	10.8	2344	9	US-09-815-242-12713	Sequence 12713, A
44	594	10.5	1831	12	US-10-282-122A-71033	Sequence 71033, A
45	554	9.8	406	16	US-10-451-467A-262	Sequence 262, App

ALIGNMENTS

RESULT 1
US-10-378-674-8
; Sequence 8, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PANTI Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATIV
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 1092
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-8

Query Match	100.0%	Score	5646	DB	15	Length	1092
Best Local Similarity	100.0%	Pred	No. 3.8e-307				
Matches	1092	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MINKKNLLTKKKPIANKSNKYAIRKFTVTGASIVIGATLLFGLGHNEAKAENSVDVK	60				
Db	1	MINKKNLLTKKKPIANKSNKYAIRKFTVTGASIVIGATLLFGLGHNEAKAENSVDVK	60				
QY	61	DSNTDDELSDSNDQSSDEEKDVNNQSDINTDDNNQIIKKETNNYDGIKESDERTES	120				
Db	61	DSNTDDELSDSNDQSSDEEKDVNNQSDINTDDNNQIIKKETNNYDGIKESDERTES	120				
QY	121	TTNVDENEATFLQKTPQDNTNLTTEBEVKSSSVSSNSSIDTAQPSHTTINREESVQTS	180				
Db	121	TTNVDENEATFLQKTPQDNTNLTTEBEVKSSSVSSNSSIDTAQPSHTTINREESVQTS	180				
QY	181	DNVEDSHVSDPANSKIKESNTSGKEENTIEOPNKVKEDSTTSQPSGYTNIDKISNQDE	240				

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 39.7921 Seconds
(without alignments)
2639.748 Million cell updates/sec

Title: US-10-806-288-15

Perfect score: 5646

Sequence: 1 MINKNNLLTKKPIANKSN.....PAGLGALLGKRRKRNKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 T30214	fibrinogen-binding
2	2698.5	47.8	1166	2 T28680	fibrinogen-binding
3	2589	45.9	1141	2 B99824	hypothetical prote
4	2234	39.6	1385	2 B99824	hypothetical prote
5	2093.5	37.1	953	2 C99824	hypothetical prote
6	1880	33.3	1315	2 T28679	fibrinogen-binding
7	1652	29.3	989	2 D99852	fibrinogen-binding
8	1628	28.8	933	2 S41539	fibrinogen-binding
9	1536	27.2	877	2 F90070	Clumping factor B
10	785	13.9	882	2 AG1671	probable peptidogl
11	782.5	13.9	903	2 AG1299	probable peptidogl
12	624.5	11.1	2271	2 F90073	hypothetical prote
13	554	9.8	406	2 S38170	SRP40 protein - ye
14	554	9.8	3394	2 T18501	hypothetical prote
15	549	9.7	2570	2 T17451	fimbriae-associate
16	525.5	9.3	1038	2 H90053	hypothetical prote
17	520	9.2	940	2 S19702	fibronectin-bindin
18	502.5	8.9	961	2 G90053	hypothetical prote
19	497.5	8.6	1018	2 A32192	fibronectin-bindin
20	471	8.3	334	2 A54138	acidic repetitive
21	471	8.3	640	2 A54502	S antigen precurs
22	452	8.0	2910	2 T28156	DNA-directed RNA p
23	442.5	7.8	1999	2 AB2018	hypothetical prote
24	435	7.7	1192	2 A71623	probable secreted
25	430.5	7.6	4550	2 T18440	hypothetical prote
26	391.5	6.9	1217	2 S52714	serpinB1 - silkw
27	388.5	6.9	695	2 S27390	calcium-binding pr
28	381.5	6.8	955	2 T18435	hypothetical prote
29	379.5	6.7	3724	2 T18427	hypothetical prote

30 371.5 6.6 1305 2 T00670 probable inositol
31 370 6.6 1063 2 D86731 hypothetical prote
32 368 6.5 1664 2 T18262 S-layer protein -
33 366 6.5 3844 2 T18402 asparagine/asparta
34 364.5 6.5 792 2 T42963 hypothetical prote
35 363.5 6.4 1360 2 T18403 asparagine/asparta
36 362 6.4 1193 2 G71605 hypothetical prote
37 351.5 6.2 1428 2 T08852 lustrin A - Calif
38 345.5 6.1 2481 2 D90011 FmB protein (limp
39 343.5 6.1 1658 2 S55101 hypothetical prote
40 338 6.0 4776 2 E95206 cell wall surface
41 336 6.0 2485 1 H71621 serine/threonine-s
42 335.5 5.9 4910 2 S64942 probable membrane
43 333.5 5.9 1072 2 A86827 hypothetical prote
44 333 5.9 797 2 A36811 hypothetical prote
45 331.5 5.9 1419 2 T30531 agglutinin-like ad

ALIGNMENTS

RESULT 1
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guse, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:98261511; PMID:9596732
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NIL>
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 5646; DB 2; Length 1092;
Best Local Similarity 100.0%; Pred. No. 3.3e-211;
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MINKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVDVK 60
Db 1 MINKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVDVK 60
Qy 61 DSNTDDELSDNDQSSDEEKNDVINNNQSNINTDNNQIIKKETNNYDGIKRSSEDRTES 120
Db 61 DSNTDDELSDNDQSSDEEKNDVINNNQSNINTDNNQIIKKETNNYDGIKRSSEDRTES 120
Qy 121 TTWVDENEATFLOKTPQDNTHLTREEVKSSSVSSSSSDTAQOPSHTTTINREESVQTS 180
Db 121 TTWVDENEATFLOKTPQDNTHLTREEVKSSSVSSSSSDTAQOPSHTTTINREESVQTS 180
Qy 181 DNVEDSHVSPFANSKIKESNTSGKENTIREQPNKVEDSTTSQPSGVNTIDKISNDE 240
Db 181 DNVEDSHVSPFANSKIKESNTSGKENTIREQPNKVEDSTTSQPSGVNTIDKISNDE 240
Qy 241 LLNLPINEYENKARPLSTTSAQPSIKKVTNQLAAEQSNVNHIIKVTQDSITEGYDDSE 300
Db 241 LLNLPINEYENKARPLSTTSAQPSIKKVTNQLAAEQSNVNHIIKVTQDSITEGYDDSE 300
Qy 301 GVTKAHDAENLIYDVTPFVDDKVKSGDTMTVDKNTVPSDLTSPFTPKIKNSGRIIA 360
Db 301 GVTKAHDAENLIYDVTPFVDDKVKSGDTMTVDKNTVPSDLTSPFTPKIKNSGRIIA 360
Qy 361 TGTVDNKNQITFTPTDYVDKYNIAKHLKLTSTVDKSKVPNNNTKLDVEYKTLSSVYK 420
Db 361 TGTVDNKNQITFTPTDYVDKYNIAKHLKLTSTVDKSKVPNNNTKLDVEYKTLSSVYK 420
Qy 421 TITVVEQRPNNENTANLQSMFTNIDTKNHTVEQTIYNPLRYSKAKETNNVNSGNGDEGST 480
Db 421 TITVVEQRPNNENTANLQSMFTNIDTKNHTVEQTIYNPLRYSKAKETNNVNSGNGDEGST 480

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:34:15 ; Search time 22.8315 Seconds
(without alignments)
2490.441 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKNNLLTKKPIANKSN.....FAGLALLGKRKRNRKN 1092

Scoring table: BLOSUM62

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1319.5	23.4	1637	1	MRSP STAAU
2	708.5	12.5	1253	1	DSPP HUMAN
3	675.5	12.0	934	1	DSPP MOUSE
4	606.5	10.7	687	1	DSPP RAT
5	554	9.8	406	1	SR40 YEAST
6	487.5	8.6	1018	1	FNBA STAAU
7	471	8.3	640	1	SANT PLAFW
8	388.5	6.9	695	1	ARP_EUGR
9	368	6.5	1664	1	SLPI_CLOTM
10	361.5	6.4	400	1	RTOA_DICDI
11	343.5	6.1	1658	1	YM67 YEAST
12	335.5	5.9	4910	1	MDN1 YEAST
13	333	5.9	797	1	VG48_HSVSA
14	331.5	5.9	1419	1	ALAI_CANAL
15	315	5.6	937	1	HYRI_CANAL
16	314	5.6	2492	1	ATRX_PANTR
17	311	5.5	2452	1	RPB1_PLAFD
18	310	5.5	1189	1	YJH6 YEAST
19	303	5.4	2492	1	ATRX_HUMAN
20	301.5	5.3	503	1	DMPI_MOUSE
21	297	5.3	2492	1	ATRX_PONFY
22	295.5	5.2	489	1	DMPI_RAT
23	294.5	5.2	429	1	DR48 YEAST
24	291	5.2	1744	1	TANA_XENLA
25	288	5.1	389	1	SERI_BOMMO
26	285.5	5.1	513	1	DMPI_HUMAN
27	283.5	5.0	1337	1	DEXT_STRDO
28	283	5.0	514	1	FBPA_BACSU
29	276.5	4.9	1085	1	IFH1 YEAST
30	276	4.9	500	1	GAR2_SCHPO
31	271.5	4.8	1025	1	MK21 YEAST
32	271	4.8	688	1	LIP_STAEP
33	271	4.8	2476	1	ATRX_MOUSE

RESULT 1

ID	MRSP STAAU	STANDARD	PRT	1637 AA
AC	P80544	O92F62		
DT	01-FEB-1995	(Rel. 33, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Methicillin-resistant surface protein precursor.			
DE	PLS.			
GN	Staphylococcus aureus.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=Isolate 1061;			
RX	MEDLINE=21189215; PubMed=11292719;			
RA	Savolainen K., Paulin L., Westerlund-Wikstrom B., Foster T.J.,			
RA	Korhonen T.K., Kuusela P.,			
RT	"Expression of p1s, a gene closely associated with the mecA gene of			
RT	methicillin-resistant Staphylococcus aureus, prevents bacterial			
RT	adhesion in vitro."			
RL	Infect. Immun. 69:3013-3020(2001).			
RN	[2]			
RP	SEQUENCE OF 624-628; 676-682; 938-948; 1156-1168; 1176-1185;			
RP	1199-1205 AND 1217-1224.			
RC	STRAIN=Isolate 1061;			
RX	MEDLINE=96270743; PubMed=8665912;			
RA	Hilden P., Savolainen K., Tynnelae J., Vuento M., Kuusela P.;			
RT	"Purification and characterisation of a plasmin-sensitive surface			
RT	protein of Staphylococcus aureus."			
RL	Eur. J. Biochem. 236:1904-910(1996).			
CC	-!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
EMBL	AF115379; AAD09131.1; -			
DR	InterPro; IPR005877; Gpos_Y5IRK.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF04650; Y5IRK_signal; 1.			
DR	TIGRFAMs; TIGR01167; LPXTG_anchor; 1.			
DR	TIGRFAMs; TIGR01168; Y5IRK_signal; 1.			
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.			
KW	Antibiotic resistance; Glycoprotein; Cell wall; Peptidoglycan-anchor;			
Repeat; Signal.	-----			
FT	CHAIN	1	48	POTENTIAL.
FT	CHAIN	49	1601	METHICILLIN-RESISTANT SURFACE PROTEIN.
FT	PROPEP	1502	1637	REMOVED BY SORTASE (POTENTIAL).
FT	DOMAIN	1301	1582	141 X 2 AA TANDEM REPEATS OF D-[SAG].

Q09624 caenorhabdi
P22290 plasmodium
Q00799 plasmodium
P87498 gallus gall
Q03400 plasmodium
P35662 bos taurus
P39521 saccharomyc
P34487 caenorhabdi
O84419 chlamydia t
P08723 rattus norv
P46949 saccharomyc
Q95120 bos taurus

ALIGNMENTS

GanCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 122.638 Seconds
(without alignments)
2809.455 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKNNLLTKKPIANKSN.....FAGLGALLGKRKKRKNKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 070022	070022 staphylococ
2	5440	96.4	1056	16 Q8CQ72	Q8CQ72 staphylococ
3	4485	79.4	931	2 Q9K113	Q9K113 staphylococ
4	2756	48.8	1171	2 Q9KXW6	Q9KXW6 staphylococ
5	2698.5	47.8	1166	2 Q86489	Q86489 staphylococ
6	2589	45.9	1141	16 Q99W46	Q99W46 staphylococ
7	2589	45.9	1141	16 Q8NXX5	Q8NXX5 staphylococ
8	2586	45.8	1141	16 Q932P7	Q932P7 staphylococ
9	2344.5	41.5	1633	16 Q8CMP4	Q8CMP4 staphylococ
10	2303.5	40.8	1733	2 Q9K114	Q9K114 staphylococ
11	2234	39.6	1385	16 Q99W47	Q99W47 staphylococ
12	2093.5	37.1	953	16 Q99W48	Q99W48 staphylococ
13	2075.5	36.8	955	16 Q8NXX7	Q8NXX7 staphylococ
14	2037	36.1	1347	16 Q8NXX6	Q8NXX6 staphylococ
15	2023.5	35.8	947	2 Q86487	Q86487 staphylococ
16	2006	35.5	1893	2 Q8KWL1	Q8KWL1 staphylococ

17	1880	33.3	1315	2	Q86488	Q86488 staphylococ
18	1652	29.3	989	16	Q99VJ4	Q99VJ4 staphylococ
19	1646.5	29.2	946	16	Q8NXX1	Q8NXX1 staphylococ
20	1628	28.8	933	2	Q53653	Q53653 staphylococ
21	1610	28.5	913	2	Q86476	Q86476 staphylococ
22	1581	28.0	935	16	Q932C5	Q932C5 staphylococ
23	1573	27.9	907	16	Q8NDL0	Q8NDL0 staphylococ
24	1536	27.2	877	16	Q99R07	Q99R07 staphylococ
25	1471	26.1	881	2	Q93MH7	Q93MH7 staphylococ
26	1216	21.5	1698	2	Q9LC00	Q9LC00 staphylococ
27	1203.5	21.3	3360	16	Q88XB6	Q88XB6 lactobacill
28	785	13.9	882	16	Q92AK9	Q92AK9 listeria in
29	782.5	13.9	903	16	Q8V697	Q8V697 listeria mo
30	747.5	13.2	2020	5	Q8ID80	Q8ID80 plasmodium
31	671.5	11.9	487	2	Q9K112	Q9K112 plasmodium
32	670.5	11.9	970	11	Q8VBY1	Q8VBY1 rattus norv
33	654.5	11.6	540	2	Q8KR21	Q8KR21 staphylococ
34	645.5	11.4	2310	16	Q8CMU7	Q8CMU7 staphylococ
35	640.5	11.3	485	16	Q8CNU7	Q8CNU7 staphylococ
36	639	11.3	579	13	Q7T138	Q7T138 brachydanio
37	624.5	11.1	2271	16	Q99QY4	Q99QY4 staphylococ
38	622	11.0	968	5	Q8ISW9	Q8ISW9 plasmodium
39	616.5	10.9	2275	16	Q8NUJ3	Q8NUJ3 staphylococ
40	612	10.8	2283	2	Q8VO99	Q8VO99 staphylococ
41	607	10.8	2900	5	Q8I2E0	Q8I2E0 plasmodium
42	565	10.0	1182	5	Q8ID30	Q8ID30 plasmodium
43	560	9.9	970	16	Q8DYL7	Q8DYL7 streptococc
44	554	9.8	3394	5	Q77384	Q77384 plasmodium
45	549	9.7	2570	2	Q9ZPF9	Q9ZPF9 streptococc

ALIGNMENTS

RESULT 1

ID	070022	PRELIMINARY;	PRT; 1092 AA.
AC	070022;		
DT	01-AUG-1998 (TrEMBLrel. 07, Created)		
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Fibrinogen-binding protein precursor.		
OS	Staphylococcus epidermidis.		
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
OX	NCBI_TaxID=1282;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HB;		
RX	MEDLINE=98261511; PubMed=9596732;		
RA	Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;		
RT	"A Fibrinogen-binding protein of Staphylococcus epidermidis."		
RL	Infect. Immun. 66:2666-2673(1998).		
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY AN AMIDE BOND (BY SIMILARITY).		
DR	EMBL; Y17116; CAA76638.1; -		
DR	PIR; T30214; T30214.		
DR	GO; GO:0005618; C:cell wall; IEA.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	InterPro; IPR008454; Cna_B.		
DR	InterPro; IPR008877; Gpos_Y5IRK.		
DR	InterPro; IPR001899; Gram_pos_anchor		
DR	Pfam; PF05738; Cna_B; 2.		
DR	Pfam; PF00746; Gram_pos_anchor; 1.		
DR	Pfam; PF04650; Y5IRK_signal; 1.		
DR	TIGRFAMs; TIGR01167; LPXTG_anchor; 1.		
DR	TIGRFAMs; TIGR01168; Y5IRK_signal; 1.		
DR	PROSITE; PS00847; GRAM_POS_ANCHORING; 1.		
KW	Cell wall; Peptidoglycan-anchor; Signal.		
FT	SIGNAL 1 51 POTENTIAL.		
CH	CHAIN 52 1092 FIBRINOGEN-BINDING PROTEIN.		
SQ	SEQUENCE 1092 AA; 119292 MW; 6542BC39AD8B984 CRC64;		
Query Match	100.0%; Score 5646; DB 2; Length 1092;		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 122.638 Seconds
(without alignments)
2515.875 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKNNLLTKKPIANKSN.....FAGLGALLGKRRKRNKNK 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 AAW41602	Aaw41602 Staphylococcus epidermidis
2	5646	100.0	1092	7 ABM79019	Abm79019 Staphylococcus epidermidis
3	4490	79.5	930	5 ABP40469	Abp40469 Staphylococcus epidermidis
4	4485	79.4	991	3 AAY83171	Aay83171 Cell wall
5	4485	79.4	991	3 AAY83171	Aay83171 Cell wall
6	4307	76.3	892	6 ABU42557	Abu42557 Protein e
7	3351.5	59.4	670	6 ABU42520	Abu42520 Protein e
8	2787	49.4	549	7 ABM79020	Abm79020 Staphylococcus epidermidis
9	2698.5	47.8	1166	2 AAY08643	Aay08643 S. aureus
10	2624.5	47.8	1166	6 ABJ18982	Abj18982 Pathogen
11	2624.5	46.5	560	7 ABM79015	Abm79015 Staphylococcus epidermidis
12	2586	45.8	1141	6 ABU42327	Abu42327 Protein e
13	2339.5	41.4	1633	6 ABU42513	Abu42513 Protein e
14	2303.5	40.8	1802	3 AAY83170	Aay83170 Cell wall
15	2303.5	40.8	1802	3 AAY83170	Aay83170 Cell wall
16	2256.5	40.0	995	6 ABM72437	Abm72437 Staphylococcus epidermidis
17	2234	39.6	1385	6 ABM72437	Abm72437 Staphylococcus epidermidis
18	2205.5	39.1	1920	6 ABU43489	Abu43489 Protein e
19	2093.5	37.1	953	6 ABU16533	Abu16533 Protein e
20	2070	36.7	1349	6 ABM72436	Abm72436 Staphylococcus epidermidis
21	2069	36.6	1349	4 AAU37544	Aau37544 Staphylococcus epidermidis
22	2069	36.6	1349	4 AAU34402	Aau34402 Staphylococcus epidermidis
23	2032.5	36.0	947	6 ABJ18940	Abj18940 Pathogen
24	2011	35.6	932	4 AAU36845	Aau36845 Staphylococcus epidermidis
25	2011	35.6	932	4 AAU34082	Aau34082 Staphylococcus epidermidis

26	1990	35.2	930	2 AAY08641	Aay08641 S. aureus
27	1880	33.3	1315	2 AAY08642	Aay08642 S. aureus
28	1880	33.3	1315	6 ABJ18969	Abj18969 Pathogen
29	1672	29.6	343	7 ABM79016	Abm79016 Staphylococcus epidermidis
30	1634	28.9	1021	4 AAU33975	Aau33975 Staphylococcus epidermidis
31	1634	28.9	1021	4 AAU33975	Aau33975 Staphylococcus epidermidis
32	1631	28.9	933	6 ABJ18947	Abj18947 Pathogen
33	1629.5	28.9	933	2 AAU89801	Aau89801 Staphylococcus epidermidis
34	1628	28.8	933	3 AAY58435	Aay58435 Staphylococcus epidermidis
35	1628	28.8	933	4 AAB69508	Aab69508 Staphylococcus epidermidis
36	1625	28.8	927	6 ABM72221	Abm72221 Staphylococcus epidermidis
37	1617.5	28.6	913	2 AAY08640	Aay08640 S. aureus
38	1611	28.5	913	6 ABJ18917	Abj18917 Pathogen
39	1581	28.0	935	6 ABU16402	Abu16402 Protein e
40	1562	27.7	316	7 ABM79017	Abm79017 Staphylococcus epidermidis
41	1552	27.5	877	6 ADA89539	Ada89539 Staphylococcus epidermidis
42	1552	27.5	877	6 ABM72702	Abm72702 Staphylococcus epidermidis
43	1536	27.2	877	6 ABU42504	Abu42504 Protein e
44	1446	25.6	287	4 AAG81687	Aag81687 S. epidermidis
45	1441.5	25.5	841	4 AAU37158	Aau37158 Staphylococcus epidermidis

ALIGNMENTS

RESULT 1
AAW41602
ID AAW41602 standard; protein; 1092 AA.
XX AC AAW41602;
XX AC

17-OCT-2003 (revised)
22-JUN-1998 (first entry)

Staphylococcus epidermidis fibrinogen binding protein FIG.

Fibrinogen binding protein; FIG; aggregation; infection;
coagulase-negative Staphylococcus; therapy; diagnosis; immunisation;
immunogen; vaccine.

Staphylococcus epidermidis; strain HB.

Key Location/Qualifiers
FT Peptide 1..51
FT Protein 52..1092
FT Region 52..824
FT /label= Sig_peptide
FT /label= Mat_protein

FT /notes= "non-repetitive region, harbours fibrinogen binding activity"

FT Region 825..1040
FT /note= "Asp-Ser dipeptide repeat region"
FT 1053..1057
FT /note= "cell wall anchoring motif"

WO9748727-A1.

24-DEC-1997.

18-JUN-1997; 97WO-SE001091.

20-JUN-1996; 96SE-00002496.

(GUSS/) GUSS B.

(NILS/) NILSSON M.

(FRYK/) FRYKBERG L.

(FLOCK/) FLOCK J.

(LIND/) LINDBERG M.

Guss B, Nilsson M, Frykberg L, Flock J, Lindberg M,

WPI, 1998-063079/06.

N-PSDB; AAV04279.

Index	Value	Label
1	3013	100.0
2	3013	100.0
3	2810.5	93.3
4	1160.5	38.5
5	1144.5	38.0
6	1143.5	38.0
7	1137.5	37.8
8	1137.5	37.8
9	576	19.1
10	567	18.8
11	559	18.6
12	499	16.6
13	497	16.5
14	489	16.2
15	485.5	16.1
16	470.5	15.6
17	470.5	15.6
18	470.5	15.6
19	470.5	15.6
20	470.5	15.6
21	470.5	15.6
22	470.5	15.6
23	470.5	15.6
24	470.5	15.6
25	470.5	15.6
26	470.5	15.6
27	470.5	15.6
28	470.5	15.6
29	470.5	15.6
30	470.5	15.6
31	470.5	15.6
32	470.5	15.6
33	470.5	15.6
34	470.5	15.6
35	470.5	15.6
36	470.5	15.6
37	470.5	15.6
38	470.5	15.6
39	470.5	15.6
40	470.5	15.6
41	470.5	15.6
42	470.5	15.6
43	470.5	15.6
44	470.5	15.6
45	470.5	15.6
46	470.5	15.6
47	470.5	15.6
48	470.5	15.6
49	470.5	15.6
50	470.5	15.6
51	470.5	15.6
52	470.5	15.6
53	470.5	15.6
54	470.5	15.6
55	470.5	15.6
56	470.5	15.6
57	470.5	15.6
58	470.5	15.6
59	470.5	15.6
60	470.5	15.6
61	470.5	15.6
62	470.5	15.6
63	470.5	15.6
64	470.5	15.6
65	470.5	15.6
66	470.5	15.6
67	470.5	15.6
68	470.5	15.6
69	470.5	15.6
70	470.5	15.6
71	470.5	15.6
72	470.5	15.6
73	470.5	15.6
74	470.5	15.6
75	470.5	15.6
76	470.5	15.6
77	470.5	15.6
78	470.5	15.6
79	470.5	15.6
80	470.5	15.6
81	470.5	15.6
82	470.5	15.6
83	470.5	15.6
84	470.5	15.6
85	470.5	15.6
86	470.5	15.6
87	470.5	15.6
88	470.5	15.6
89	470.5	15.6
90	470.5	15.6
91	470.5	15.6
92	470.5	15.6
93	470.5	15.6
94	470.5	15.6
95	470.5	15.6
96	470.5	15.6
97	470.5	15.6
98	470.5	15.6
99	470.5	15.6
100	470.5	15.6

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:34:15 ; Search time 12.1685 Seconds
(without alignments)
2490.441 Million cell updates/sec

Title: US-10-806-288-13

Perfect score: 3013

Sequence: 1 SSDDEKNDVINNQSIINTDD.....PDGTSKSVRTDEGKYQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141691 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300.5	10.0	1018	1	FNBA STAAU
2	203	6.7	537	1	ARP FLAPA
3	202	6.7	1093	1	PIK DICDI
4	196.5	6.5	1139	1	HMW1 MYCGE
5	192.5	6.4	2867	1	BBP2 PLAVE
6	192	6.4	918	1	YMBJ CAEHL
7	189.5	6.3	1664	1	INT1 CANAL
8	188.5	6.3	578	1	LIPA WYCPU
9	188.5	6.3	1435	1	ERAI PLAFU
10	186.5	6.2	3110	1	LM22 HUMAN
11	184.5	6.1	1744	1	TANA XENLA
12	184	6.1	817	1	YG4A YEAST
13	183.5	6.1	2009	1	SECT YEAST
14	183	6.1	2452	1	RPB1 PLAFD
15	182.5	6.1	1096	1	MRC1 YEAST
16	182.5	6.1	1381	1	YBET YEAST
17	182.5	6.1	1658	1	YBET YEAST
18	182	6.0	1419	1	ALAI CANAL
19	182	6.0	1460	1	N159 YEAST
20	181.5	6.0	1093	1	YKDS CAEHL
21	179.5	6.0	1165	1	YNF4 YEAST
22	179	5.9	1253	1	DSPH HUMAN
23	178	5.9	960	1	YMX6 YEAST
24	178	5.9	1875	1	MLP1 YEAST
25	177	5.9	2660	1	YEEJ BCO57
26	176	5.8	1196	1	BXCN GLOB
27	175.5	5.8	1183	1	CNA STAAU
28	175.5	5.8	1637	1	MRSP STAAU
29	175	5.8	1147	1	CGAI HELPY
30	174	5.8	2022	1	ANT1 ONCVO
31	173.5	5.8	967	1	Y5G0 GLOB
32	173	5.7	1310	1	YB35 SCHPO
33	172	5.7	682	1	NISP_LACLA

ALIGNMENTS

RESULT 1

ID	FNBA STAAU	STANDARD;	PRT;	1018 AA.
AC	P14738,			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Fibronectin-binding protein precursor (FNBP).			
GN	FNBA			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NTCT 8325-4;			
RA	MEDLINE=8909898; PubMed=2521391;			
RA	Signaes C., Raucii G., Joensson K., Lindgren P.-E.,			
RA	Anantharamaiah G.M., Hoeoek M., Lindberg M.;			
RT	"Nucleotide sequence of the gene for a fibronectin-binding protein			
RT	from Staphylococcus aureus: use of this peptide sequence in the			
RT	synthesis of biologically active peptides."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989)			
CC	- - FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN			
CC	PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE			
CC	WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO			
CC	THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S.AUREUS,			
CC	THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE			
CC	INVASION.			
CC	- - SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.iab-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
CC	EMBL; J04151; AAA26632.1; -			
DR	InterPro; IPR004237; Fn bind.			
DR	InterPro; IPR005877; Gpos_YSRK.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Pfam; PF02986; Fn bind; 1.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF04650; YSRK signal; 1.			
DR	TIGRFAMs; TIGR01167; LPKXG anchor; 1.			
DR	TIGRFAMs; TIGR01168; YSRK signal; 1.			
DR	PROSITE; PS00847; GRAM_POS_ANCHORING; 1.			
KW	Cell wall; Peptidoglycan-anchor; Repeat; Signal.			
FT	SIGNAL 1 36			
FT	CHAIN 37 985			
FT	PROPEP 986 1018			
FT	REPEAT 545 574			
FT	REPEAT 575 604			
FT	DOMAIN 745 878			

Q02510 staphylococ
P53550 saccharomyc
P46590 candida alb
P07866 saccharomyc
P25558 saccharomyc
P47024 saccharomyc
P21131 streptococc
P25062 halobacteri
Q02192 streptococc
Q00798 plasmodium
P14164 saccharomyc
P47134 saccharomyc

34 171.5 5.7 688 1 LIP STARP
35 170.5 5.7 970 1 PSUI YEAST
36 170.5 5.7 1260 1 ALSI CANAL
37 170.5 5.7 1435 1 LTR1 YEAST
38 170.5 5.7 1636 1 BUD3 YEAST
39 170 5.6 1803 1 YUL3 YEAST
40 169.5 5.6 571 1 TACV STRPY
41 169.5 5.6 827 1 CSG HALVO
42 169.5 5.6 1020 1 BCA STRAG
43 169.5 5.6 2869 1 RBF1 PLAVB
44 169 5.6 731 1 BAF1 YEAST
45 169 5.6 954 1 BIR1 YEAST

4 X APPROXIMATE TANDEM REPEATS,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 21.2079 Seconds
(without alignments)
2639.748 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEKNDVNNNQSIINTDD.....PDGTSKSVRTDEGKVPDQ 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3013	100.0	1092	2 T30214	fibrinogen-binding
2	1144.5	38.0	1166	2 T28680	fibrinogen-binding
3	1137.5	37.8	1141	2 E89824	hypothetical prote
4	576	19.1	1315	2 T28679	fibrinogen-binding
5	559	18.6	1385	2 D89824	hypothetical prote
6	489	16.2	953	2 C89824	hypothetical prote
7	452	15.0	933	2 S41539	fibrinogen-binding
8	452	15.0	989	2 D89852	fibrinogen-binding
9	397.5	13.2	877	2 F90070	Clumping factor B
10	376	12.5	940	2 S19702	fibrinectin-bindin
11	344.5	11.4	961	2 G90053	hypothetical prote
12	326	10.8	1038	2 H90053	hypothetical prote
13	300.5	10.0	1018	2 A32192	fibrinectin-bindin
14	239	7.9	1039	2 T30856	protein F2 - Strept
15	225	7.5	1463	2 T30290	AAS surface protei
16	216.5	7.2	1959	2 AG1085	hypothetical prote
17	216.5	7.2	3394	2 T18501	hypothetical prote
18	212.5	7.1	1125	2 G90598	membrane nuclease,
19	211.5	7.0	2206	2 T71611	hypothetical prote
20	208.5	6.9	1072	2 A86827	hypothetical prote
21	208	6.9	1127	2 T28317	ORF MSV156 bypothe
22	208	6.9	3724	2 T18427	hypothetical prote
23	207.5	6.9	1711	2 T18429	hypothetical prote
24	207.5	6.9	2401	2 T28676	rhoptry protein -
25	207	6.9	769	2 F89870	serine proteinase
26	207	6.9	1939	2 T18372	repeat organellar
27	206.5	6.9	4550	2 T18440	hypothetical prote
28	206.5	6.9	821	2 S67087	hypothetical prote
29	206	6.8	1302	1 JC6009	surface-located me

RESULT 1

T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C/Accession: T30214

R/Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A/Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A/Reference number: Z20781; MUID:98261511; PMID:9596732

A/Accession: T30214
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1092 <NIL>

A/Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 3013; DB 2; Length 1092;
Best Local Similarity: 100.0%; Pred. No. 7.2e-123;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

30	205.5	6.8	1365	2 T30822	lmp1 protein - Myc
31	205.5	6.8	2510	2 T28160	hypothetical prote
32	205	6.8	2481	2 D90011	FmtB protein [lmpo
33	204	6.8	3844	2 T18402	asparagine/asparta
34	203	6.7	537	2 A23770	asparagine-rich pr
35	202.5	6.7	1997	2 T71607	DNA helicase II BR
36	202.5	6.7	2369	2 T28677	rhoptry protein -
37	202	6.7	1093	2 T18275	1-phosphatidylinos
38	201.5	6.7	4688	2 F82885	hypothetical prote
39	199	6.6	665	2 B71609	hypothetical prote
40	198.5	6.6	807	2 B71605	hypothetical prote
41	198.5	6.6	1805	2 T02712	similar to late em
42	198.5	6.6	3147	2 T18674	hypothetical prote
43	198	6.6	1308	2 B71622	probable membrane
44	197	6.5	940	2 AD1374	internalin protein
45	197	6.5	2523	2 T18477	hypothetical prote

ALIGNMENTS

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 68.491 Seconds
(without alignments)
2734.476 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEKNDVINNQSIINTDD.....PDGTSKSVRTDEGKYQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 31799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	15	US-10-378-674-8
2	2810.5	93.3	892	12	US-10-282-122A-70481
3	2810.5	93.3	892	16	US-10-661-809-21
4	2810.5	93.3	930	12	US-10-615-383-10
5	2810.5	93.3	930	16	US-10-689-082-10
6	2810.5	93.3	930	16	US-10-690-184-10
7	2670	88.6	549	15	US-10-378-674-9
8	2510.5	83.3	560	15	US-10-378-674-2
9	1672	55.5	343	15	US-10-378-674-4
10	1562	51.8	316	15	US-10-378-674-6
11	1137.5	37.8	1141	12	US-10-282-122A-70251
12	1116.5	37.1	670	12	US-10-282-122A-70444
13	576	19.1	1349	9	US-09-815-242-5898
14	576	19.1	1349	9	US-09-815-242-13137
15	559	18.6	1385	12	US-10-282-122A-44324

16	499.5	16.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
17	499.5	16.6	932	9	US-09-815-242-12438	Sequence 12438, A
18	499	16.6	1742	12	US-10-615-383-4	Sequence 4, Appl1
19	499	16.6	1742	16	US-10-689-082-4	Sequence 4, Appl1
20	499	16.6	1742	16	US-10-690-184-4	Sequence 4, Appl1
21	493	16.4	1633	12	US-10-282-122A-70437	Sequence 70437, A
22	489	16.2	953	12	US-10-282-122A-44457	Sequence 44457, A
23	486.5	16.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
24	455	15.1	841	9	US-09-815-242-5779	Sequence 5779, Ap
25	455	15.1	841	9	US-09-815-242-12751	Sequence 12751, A
26	454	15.1	936	8	US-08-781-986A-5249	Sequence 5249, Ap
27	454	15.1	936	12	US-10-329-624-5249	Sequence 5249, Ap
28	452	15.0	935	12	US-10-282-122A-44326	Sequence 44326, A
29	449	14.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
30	449	14.9	1021	9	US-09-815-242-12544	Sequence 12544, A
31	443.5	14.7	520	14	US-10-056-052-2	Sequence 2, Appl1
32	397.5	13.2	877	12	US-10-282-122A-70428	Sequence 70428, A
33	377	12.5	496	14	US-10-311-879-29	Sequence 29, Appl1
34	376	12.5	767	9	US-09-815-242-5899	Sequence 5899, Ap
35	376	12.5	767	9	US-09-815-242-13140	Sequence 13140, A
36	349	11.6	331	14	US-10-056-052-4	Sequence 4, Appl1
37	349	11.6	345	9	US-09-813-820-7	Sequence 7, Appl1
38	344.5	11.4	961	12	US-10-282-122A-43778	Sequence 43778, A
39	326	10.8	1038	12	US-10-282-122A-43827	Sequence 43827, A
40	319.5	10.6	978	9	US-09-815-242-5456	Sequence 5456, Ap
41	319.5	10.6	1001	9	US-09-815-242-12686	Sequence 12686, A
42	308.5	10.2	1018	9	US-09-815-242-5797	Sequence 5797, Ap
43	308.5	10.2	1018	9	US-09-815-242-12838	Sequence 12838, A
44	300.5	10.0	1027	8	US-08-781-986A-5254	Sequence 5254, Ap
45	300.5	10.0	1027	12	US-10-329-624-5254	Sequence 5254, Ap

ALIGNMENTS

RESULT 1

US-10-378-674-8
; Sequence 8, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PARTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT
; FILE OF INVENTION: STAPHYLOCOCCAL PROTEINS
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 1092
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-8

Query Match	100.0%;	Score 3013;	DB 15;	Length 1092;
Best Local Similarity	100.0%;	Pred. No. 1.3e-176;		
Matches 582;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	SSDEKNDVINNQSIINTDDNNQIIKKEETNNYDGIKESDRTSTTNNVDENEATFLOK 60		
DB	75	SSDEKNDVINNQSIINTDDNNQIIKKEETNNYDGIKESDRTSTTNNVDENEATFLOK 134		
QY	61	TPQDNTHLTEREERKSSSVESNSSIDTAQPSHTTINREESVQTSNDVDSHVDFANS 120		
DB	135	TPQDNTHLTEREERKSSSVESNSSIDTAQPSHTTINREESVQTSNDVDSHVDFANS 194		
QY	121	KIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELLNLPINEYENKAR 180		
DB	195	KIKESNTSGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELLNLPINEYENKAR 254		
QY	181	PLSTTSQAQPSIKRVTNNQLAEOGQNVNHLIKVTOQSTEGYDDSEGVKAHAENLIYD 240		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 17.0358 Seconds
(without alignments)
1763.712 Million cell updates/sec

Title: US-10-806-288-13

Perfect score: 3013

Sequence: 1 SSDEKNDVINNQSIINTDD.....PDGTSKSVRTDCKGQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgm2_6/ptodata/2/iaa/5A.COMB.pdp.*
- 2: /cgm2_6/ptodata/2/iaa/5B.COMB.pdp.*
- 3: /cgm2_6/ptodata/2/iaa/6A.COMB.pdp.*
- 4: /cgm2_6/ptodata/2/iaa/6B.COMB.pdp.*
- 5: /cgm2_6/ptodata/2/iaa/PCTUS.COMB.pdp.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2810.5	93.3	930	4	US-09-134-001C-5314
2	2810.5	93.3	930	4	US-09-386-962C-10
3	1144.5	38.0	1166	4	US-09-200-650B-7
4	576	19.1	1315	4	US-09-200-650B-5
5	499	16.6	1742	4	US-09-386-962C-4
6	455	15.1	930	4	US-09-200-650B-3
7	454	15.1	936	4	US-08-956-171E-5249
8	452	15.0	933	3	US-08-293-728-2
9	452	15.0	933	3	US-09-421-868-2
10	396	13.1	918	4	US-09-200-650B-1
11	349	11.6	345	3	US-08-856-253-7
12	300.5	10.0	1027	4	US-08-956-171E-5254
13	259.5	8.6	1112	2	US-08-714-402-2
14	259.5	8.6	1161	4	US-09-327-536-2
15	220.5	7.3	3696	4	US-09-134-001C-5080
16	209	6.9	630	3	US-08-973-462-9
17	208.5	6.9	251	4	US-08-956-171E-5252
18	203.5	6.8	2504	4	US-09-328-352-5921
19	202.5	6.7	368	4	US-09-134-000C-4597
20	202	6.7	676	4	US-09-134-001C-4318
21	201.5	6.7	1177	4	US-09-134-001C-5106
22	197	6.5	2314	4	US-09-268-347-49
23	195	6.5	1335	4	US-09-134-001C-3716
24	195	6.5	10182	4	US-09-134-001C-3159
25	192.5	6.4	1306	4	US-09-134-000C-6670
26	192.5	6.4	1786	3	US-08-973-462-8
27	190.5	6.3	1833	4	US-08-621-944A-4

28	190.5	6.3	1833	4	US-08-945-567D-4	Sequence 4, Appl1
29	190.5	6.3	1992	4	US-08-621-944A-3	Sequence 3, Appl1
30	190.5	6.3	1992	4	US-08-945-567D-3	Sequence 3, Appl1
31	189.5	6.3	1664	1	US-09-599-652-2	Sequence 2, Appl1
32	189.5	6.3	1664	2	US-08-642-846-2	Sequence 2, Appl1
33	189.5	6.3	1664	4	US-08-642-846-2	Sequence 2, Appl1
34	188.5	6.3	1435	2	US-08-568-459A-4	Sequence 4, Appl1
35	188.5	6.3	1435	2	US-08-487-826B-4	Sequence 4, Appl1
36	188.5	6.3	1435	4	US-09-210-288-4	Sequence 4, Appl1
37	187	6.2	1416	4	US-09-071-035-404	Sequence 404, App
38	187	6.2	1448	4	US-09-071-035-402	Sequence 402, App
39	186.5	6.2	3088	4	US-09-562-702A-8	Sequence 8, Appl1
40	186.5	6.2	3089	4	US-09-562-702A-4	Sequence 4, Appl1
41	186.5	6.2	3110	4	US-09-562-702A-2	Sequence 2, Appl1
42	186.5	6.2	3110	4	US-09-562-702A-6	Sequence 6, Appl1
43	186.5	6.2	3110	4	US-09-561-709B-7	Sequence 7, Appl1
44	186.5	6.2	3111	2	US-08-460-309-4	Sequence 4, Appl1
45	186.5	6.2	3111	2	US-08-125-077-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1

US-09-134-001C-5314

; Sequence 5314, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GPC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5314

; LENGTH: 930

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5314

Query Match 93.3%; Score 2810.5; DB 4; Length 930;

Best Local Similarity 93.6%; Pred. No. 2e-158;

Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVINNQSIINTDDNNQIIKKEETNNYDIEKRSSEDRTESTTNDENEATFLQK 60

Db 73 SSNEEKNDVINNQSIINTDDNQ-IKKEETNSDAIENRSKDIQTSTTNDENEATFLQK 131

Qy 61 TPQNTLHTEEEVKSSSVSSSSSIDTAQPSHTTINREESVQTSNDVDSHVSDFPANS 120

Db 132 TPQNTQLKEEVKPSVSSSSSMVTAQPSHTTINSEASIQTSNDNEKSRVSDPANS 191

Qy 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELNLPINEYENKAR 180

Db 192 KIIESTNESKEENTIEQPNKVRSDTSQPSYKYNIDEKISNQDELNLPINEYENKVR 251

Qy 181 PLSTTSQPSIKRVTVNQLAAEQGSNNVHLIKVTDQISITEGYDSDSEGVIKAHDAENLIYD 240

Db 252 PLSTTSQPSIKRVTVNQLAAEQGSNNVHLIKVTDQISITEGYDSDSDGIIKAHDAENLIYD 311

Qy 241 VTFEVDKVKSGDVTMTVDIDKNTVPSDLTDSFTPKIKNSGEIATGTYDNKNQIYTT 300

Db 312 VTFEVDKVKSGDVTMTVDIDKNTVPSDLTDSFALPKIKNSGEIATGTYDNKNQIYTT 371

Qy 301 FTDVVDKYENIKAHKLKLTSTYIDSKVPNNNTKLDVEYKNTALSSVNNKTTIYVEYQRPENRT 360

Db 372 FTDVVDKYENIKAHKLKLTSTYIDSKVPNNNTKLDVEYKNTALSSVNNKTTIYVEYQRPENRT 431

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 65.362 Seconds
(without alignments)
2515.875 Million cell, updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEKNDVINNQSIINTDD.....PDGTSKSVRTDEGKQFQDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the chance being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 AAW41602	AAW41602 Staphyloc
2	3013	100.0	1092	7 ABM79019	Abm79019 Staphyloc
3	2810.5	93.3	892	6 ABU42557	Abu42557 Protein e
4	2810.5	93.3	930	5 ABP40469	Abp40469 Staphyloc
5	2810.5	93.3	991	3 AAY83171	Aay83171 Cell wall
6	2810.5	93.3	991	3 AAY70120	Aay70120 Staph. ep
7	2670	88.6	549	7 ABM79020	Abm79020 Staphyloc
8	2510.5	83.3	560	7 ABM79015	Abm79015 Staphyloc
9	1672	55.5	343	7 ABM79016	Abm79016 Staphyloc
10	1562	51.8	316	7 ABM79017	Abm79017 Staphyloc
11	1144.5	38.0	1166	6 AAY08643	Aay08643 S. aureus
12	1142.5	37.9	1166	6 ABJ18982	Abj18982 Pathogen
13	1137.5	37.8	1141	6 ABU42327	Abu42327 Protein e
14	1116.5	37.1	670	6 ABU42520	Abu42520 Protein e
15	857	28.4	278	4 AAG82803	Aag82803 S. epider
16	576	19.1	1315	6 AAY08642	Aay08642 S. aureus
17	576	19.1	1315	6 ABJ18969	Abj18969 Pathogen
18	576	19.1	1349	4 AAU37544	Aau37544 Staphyloc
19	576	19.1	1349	4 AAU34402	Aau34402 Staphyloc
20	576	19.1	1349	6 ABM72436	Abm72436 Staphyloc
21	559	18.6	1385	6 ABU16400	Abu16400 Protein e
22	499.5	16.6	932	4 AAU36845	Aau36845 Staphyloc
23	499.5	16.6	932	4 AAU34082	Aau34082 Staphyloc
24	499	16.6	1802	3 AAY83170	Aay83170 Cell wall
25	499	16.6	1802	3 AAY70119	Aay70119 Staph. ep

ALIGNMENTS

RESULT 1
AAW41602

ID AAW41602 standard; protein; 1092 AA.

XX AAW41602;

XX 17-OCT-2003 (revised)

DT 22-JUN-1998 (first entry)

XX Staphylococcus epidermidis fibrinogen binding protein FIG.

KW Fibrinogen binding protein; FIG; aggregation; infection;

KW coagulase-negative Staphylococcus; therapy; diagnosis; immunisation;

KW immunogen; vaccine.

OS Staphylococcus epidermidis; strain HB.

XX FH Key Location/Qualifiers

FT Peptide 1..51

FT Protein /label= Sig_peptide

FT Region /label= Mat_protein

FT Region 52..824

FT Region /note= "non-repetitive region, harbours fibrinogen binding activity"

FT Region 825..1040

FT Region /note= "Asp-Ser dipeptide repeat region"

FT Region 1053..1057

FT Region /note= "cell wall anchoring motif"

XX WO9748727-A1.

XX 24-DEC-1997.

XX 18-JUN-1997; 97WO-SE001091.

XX 20-JUN-1996; 96SE-00002496.

XX (GUSS//) GUSS B.

XX (NILS//) NILSSON M.

XX (FRYK//) FRYKBERG L.

XX (FLOC//) FLOCK J.

XX (LIND//) LINDBERG M.

XX Guss B, Nilsson M, Frykberg L, Flock J, Lindberg M;

XX WPI; 1998-063079/06.

XX N-FSDB; AAV04279.

AAg82343 S. epider
ABu42513 Protein e
ABu16533 Protein e
ABu43489 Protein e
AAu37158 Staphyloc
AAu34283 Staphyloc
Aay08641 S. aureus
ABj18947 Pathogen
ABj18940 Pathogen
ABm72437 Staphyloc
ABm72221 Staphyloc
AAy89801 Staphyloc
Aay58435 Staphyloc
Aab69508 Staphyloc
ABu16402 Protein e
AAu33975 Staphyloc
AAu36951 Staphyloc
AAe29262 Staphyloc
Ada89555 Staphyloc
Ada89539 Staphyloc

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 65.362 Seconds
(without alignments)
2515.875 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEEKNDVNNQNSINTDD.....PDGTSKSVRTDEGKQPDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003s:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 AAW41602	Aaw41602 Staphyloc
2	3013	100.0	1092	7 ABM79019	Abm79019 Staphyloc
3	2810.5	93.3	892	6 ABU42557	Abu42557 Protein e
4	2810.5	93.3	930	5 ABP40469	Abp40469 Staphyloc
5	2810.5	93.3	991	3 AAY83171	Aay83171 Cell wall
6	2810.5	93.3	991	3 AAY70120	Aay70120 Staph. ep
7	2670	88.6	549	7 ABM79020	Abm79020 Staphyloc
8	2510.5	83.3	560	7 ABM79015	Abm79015 Staphyloc
9	1672	55.5	343	7 ABM79016	Abm79016 Staphyloc
10	1562	51.8	316	7 ABM79017	Abm79017 Staphyloc
11	1144.5	38.0	1166	2 AAY08643	Aay08643 S. aureus
12	1142.5	37.9	1166	6 ABJ18982	Abj18982 Pathogen e
13	1137.5	37.8	1141	6 ABU42327	Abu42327 Protein e
14	1116.5	37.1	670	6 ABU42520	Abu42520 Protein e
15	857	28.4	278	4 AAG82803	Aag82803 S. epider
16	576	19.1	1315	2 AAY08642	Aay08642 S. aureus
17	576	19.1	1315	6 ABJ18969	Abj18969 Pathogen
18	576	19.1	1349	4 AAU37544	Aau37544 Staphyloc
19	576	19.1	1349	4 AAU34402	Aau34402 Staphyloc
20	576	19.1	1349	6 ABM72436	Abm72436 Staphyloc
21	559	18.6	1385	6 ABU16400	Abu16400 Protein e
22	499.5	16.6	932	4 AAU36845	Aau36845 Staphyloc
23	499.5	16.6	932	4 AAU34082	Aau34082 Staphyloc
24	499	16.6	1802	3 AAY83170	Aay83170 Cell wall
25	499	16.6	1802	3 AAY70119	Aay70119 Staph. ep

26	497	16.5	1155	4 AAG82343	Aag82343 S. epider
27	493	16.4	1633	6 ABU42513	Abu42513 Protein e
28	489	16.2	953	6 ABU16533	Abu16533 Protein e
29	486.5	16.1	1920	6 ABU43489	Abu43489 Protein e
30	455	15.1	841	4 AAU37158	Aau37158 Staphyloc
31	455	15.1	841	4 AAU34283	Aau34283 Staphyloc
32	455	15.1	930	2 AAY08641	Aay08641 S. aureus
33	455	15.1	933	6 ABJ18947	Abj18947 Pathogen
34	455	15.1	947	6 ABM18940	Abm18940 Pathogen
35	455	15.1	995	6 ABM72437	Abm72437 Staphyloc
36	454	15.1	927	6 ABM72221	Abm72221 Staphyloc
37	454	15.1	936	2 AAW89801	Aaw89801 Staphyloc
38	452	15.0	933	3 AAY58435	Aay58435 Staphyloc
39	452	15.0	933	4 AAB89508	Aab89508 Staphyloc
40	452	15.0	935	6 ABU16402	Abu16402 Protein e
41	449	14.9	1021	4 AAU33975	Aau33975 Staphyloc
42	449	14.9	1021	4 AAU36951	Aau36951 Staphyloc
43	443.5	14.7	520	5 AAE29262	Aae29262 Staphyloc
44	397.5	13.2	567	6 ADA89555	Ada89555 Staphyloc
45	397.5	13.2	877	6 ADA89539	Ada89539 Staphyloc

ALIGNMENTS

RESULT 1
AAW41602
ID AAW41602 standard; protein; 1092 AA.
XX AAW41602;
AC AAW41602;
XX 17-OCT-2003 (revised)
DT 22-JUN-1998 (first entry)
XX
DE Staphylococcus epidermidis fibrinogen binding protein FIG.
XX
KW Fibrinogen binding protein; FIG; aggregation; infection;
KW coagulase-negative staphylococcus; therapy; diagnosis; immunisation;
KW immunogen; vaccine.
XX
OS Staphylococcus epidermidis; strain HB.
XX
FH Key Location/Qualifiers
FT Peptide 1..51
FT Protein /label= Sig_peptide
FT Protein /label= Mat_protein
FT Region 52..824
FT /note= "non-repetitive region, harbours fibrinogen binding activity"
FT Region 825..1040
FT /note= "Asp-Ser dipeptide repeat region"
FT Region 1053..1057
FT /note= "cell wall anchoring motif"
XX
WO9748727-A1.
PN
XX
PD 24-DEC-1997.
XX
PF 18-JUN-1997; 97WO-SE001091.
XX
PR 20-JUN-1996; 96SE-00002496.
XX
PA (GUSS/) GUSS B.
PA (NILS/) NILSSON M.
PA (FRYK/) FRYKBERG L.
PA (FLOC/) FLOCK J.
PA (LIND/) LINDBERG M.
PI Guss B, Nilsson M, Frykberg L, Flock J, Lindberg M;
XX
XX WPI: 1998-063079/06.
DR N-PSDB; AAV04273.
DR

XX Fibrinogen-binding protein from coagulase-negative Staphylococcus - used
 PT for prevention, treatment and diagnosis of Staphylococcus infection.
 PS Example 3; Fig 6; 45pp; English.
 XX
 CC The protein comprises the fibrinogen binding protein (FIG) of coagulase-
 CC negative Staphylococcus epidermidis HB. Its amino acid sequence was
 CC deduced from the isolated fig gene (see AAV04279). The closest known
 CC analogue of FIG is the clumping factor of *S. aureus* which also binds
 CC fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG
 CC polypeptides can be expressed in host cells. They are used as immunogens,
 CC particularly in vaccines (which may be expressed in vivo) to protect
 CC humans and animals against coagulase-negative Staphylococcus infection.
 CC Antibodies raised against FIG can be used for passive immunisation. They
 CC block the adherence of bacteria) and for diagnosis. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 1092 AA;

Query Match 100.0%; Score 3013; DB 2; Length 1092;
 Best Local Similarity 100.0%; Pred. No. 2.5e-156;
 Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSDEKNDVINNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENAEATFLOK 60
 Db 75 SSDEKNDVINNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENAEATFLOK 134
 QY 61 TPQDNTHLTHEEVKESSESSVSIDTAQOPSHHTINREESVQTSNVEDSHVDFANS 120
 Db 135 TPQDNTHLTHEEVKESSESSVSIDTAQOPSHHTINREESVQTSNVEDSHVDFANS 194
 QY 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 180
 Db 195 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 254
 QY 181 PLSTTSAQPSIKRVTVNQLAAEOGSNNHLIKVTDQSITEGYDDSEGVKAHDAENLIYD 240
 Db 255 PLSTTSAQPSIKRVTVNQLAAEOGSNNHLIKVTDQSITEGYDDSEGVKAHDAENLIYD 314
 QY 241 VTFEVDKVKSGDWTVDIDKNTVPSDLTDSFTIPKIKNSGEIIATGYDNKNKQIYTT 300
 Db 315 VTFEVDKVKSGDWTVDIDKNTVPSDLTDSFTIPKIKNSGEIIATGYDNKNKQIYTT 374
 QY 301 FTDYVDKYENIKAHKLKLTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 360
 Db 375 FTDYVDKYENIKAHKLKLTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 434
 QY 361 ANLOSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVKYKVG 420
 Db 435 ANLOSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVKYKVG 494
 QY 421 DNQNLPSNRNIYDYSEYEDVTNDYDVAQLGNNDVNIINFGNIDSPYIIKVISKYPNKKDDY 480
 Db 495 DNQNLPSNRNIYDYSEYEDVTNDYDVAQLGNNDVNIINFGNIDSPYIIKVISKYPNKKDDY 554
 QY 481 TTIOQTVMQTTINEYTGFEPTASNDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVK 540
 Db 555 TTIOQTVMQTTINEYTGFEPTASNDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVK 614
 QY 541 DGIQNTNDNEKPLSNVLVTLTPDGTSKSVRTDDEGKYQFDG 582
 Db 615 DGIQNTNDNEKPLSNVLVTLTPDGTSKSVRTDDEGKYQFDG 656

RESULT 2

ABM79019

ID ABM79019 standard; protein; 1092 AA.

XX

AC ABM79019;

XX

DT 15-JAN-2004 (first entry)

XX

DE Staphylococcus epidermidis polypeptide.
 XX Infection; antibacterial; vaccine.
 XX

OS Staphylococcus epidermidis.
 XX

PN WO2003076470-A1.
 XX

PD 18-SEP-2003.
 XX

PF 05-MAR-2003; 2003WO-US006415.
 XX

PR 05-MAR-2002; 2002US-0361324P.
 XX

XX (INH1-) INHIBITEX INC.
 PA (TEXA) UNIV TEXAS A & M SYSTEM.
 PA

XX Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
 PI Robbins J, Vernachio J, Bowden MG;
 PI

XX WPI; 2003-722324/68.
 XX

XX New antibody recognizing a Staphylococcus epidermidis protein comprising
 PT SdrG NIN2N3, SdrG N2N3 or SdrGTR2 useful for preparing a composition for
 PT treating or preventing a coagulase-negative Staphylococcus infection.
 PT

XX Claim 16; Page 36-37; 78pp; English.
 PS

XX The present sequence comprises the protein sequence of a polypeptide of a
 CC coagulase-negative Staphylococcus epidermidis. A claimed monoclonal
 CC antibody recognises this protein and is used in a claimed method of
 CC treating or preventing a coagulase-negative staphylococcal infection in a
 CC human or animal, e.g. a nosocomial coagulase-negative staphylococcal
 CC infection in low birth weight infants
 CC

XX Sequence 1092 AA;
 SQ

Query Match 100.0%; Score 3013; DB 7; Length 1092;
 Best Local Similarity 100.0%; Pred. No. 2.5e-156;
 Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSDEKNDVINNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENAEATFLOK 60
 Db 75 SSDEKNDVINNQSDINTDDNNQIIKKEETNNYDGIKESDRTSTNNVDENAEATFLOK 134
 QY 61 TPQDNTHLTHEEVKESSESSVSIDTAQOPSHHTINREESVQTSNVEDSHVDFANS 120
 Db 135 TPQDNTHLTHEEVKESSESSVSIDTAQOPSHHTINREESVQTSNVEDSHVDFANS 194
 QY 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 180
 Db 195 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELNLNLPINEYENKAR 254
 QY 181 PLSTTSAQPSIKRVTVNQLAAEOGSNNHLIKVTDQSITEGYDDSEGVKAHDAENLIYD 240
 Db 255 PLSTTSAQPSIKRVTVNQLAAEOGSNNHLIKVTDQSITEGYDDSEGVKAHDAENLIYD 314
 QY 241 VTFEVDKVKSGDWTVDIDKNTVPSDLTDSFTIPKIKNSGEIIATGYDNKNKQIYTT 300
 Db 315 VTFEVDKVKSGDWTVDIDKNTVPSDLTDSFTIPKIKNSGEIIATGYDNKNKQIYTT 374
 QY 301 FTDYVDKYENIKAHKLKLTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 360
 Db 375 FTDYVDKYENIKAHKLKLTSYIDKSKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRT 434
 QY 361 ANLOSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVKYKVG 420
 Db 435 ANLOSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVKYKVG 494
 QY 421 DNQNLPSNRNIYDYSEYEDVTNDYDVAQLGNNDVNIINFGNIDSPYIIKVISKYPNKKDDY 480
 Db 495 DNQNLPSNRNIYDYSEYEDVTNDYDVAQLGNNDVNIINFGNIDSPYIIKVISKYPNKKDDY 554

XX	13-AUG-1998;	98US-00134001.
PF		
XX		
XX	14-AUG-1997;	97US-0055779P.
PR		
PR	08-NOV-1997;	97US-0064964P.
XX		
XX	(GENO-) GENOME THERAPEUTICS CORP.	
PA		
XX		
XX	Doucette-Stamm LA, Bush D;	
PI		
XX		
DR	WPI; 2002-381255/41.	
DR	N-PSDB; ABN93014.	
XX		
XX	Novel isolated nucleic acid encoding a Staphylococcus epidermis	
PT	polypeptide, useful for diagnosing and treating bacterial infections.	
PT		
XX		
XX	Disclosure; SEQ ID NO 5314; 267pp; English.	
PS		
XX		
CC	ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading	
CC	frame (ORF) nucleic acid sequences which encode the amino acid sequences	
CC	given in ABP35124 to ABP37960. The S. epidermidis sequences have	
CC	antibacterial activity and can be used in gene therapy. The sequences can	
CC	also be used in the diagnosis and treatment of bacterial infections,	
CC	particularly S. epidermidis infections. The sequences can be used to	
CC	screen for compounds able to interfere with the S. epidermidis life cycle	
CC	or inhibit S. epidermidis infection. N.B. The sequence data for this	
CC	patent did not form part of the printed specification, but was obtained	
CC	in electronic format directly from the USPTO web site	
XX		
SQ	Sequence 930 AA;	

[illegible]

612 DGIQNTWDNKPLSNVLVTLYTPDGTSGSVRTDEEGKQFDG 653

```

Query Match          93.3%; Score 2810.5; DB 3; Length 991;
Best Local Similarity 93.6%; Pred. No. 2.6e-145;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY      1  SDEEKNDVINNNOSINTDDNQIIKKEETNNYDGIKRSDEKSTNTVDNEATFLQK 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

QY 421 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480
DB 525 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 584
QY 481 TTIOQTVMOTTINEYTGERTASYDNTIAFSTSSGOGGDLPPKTYKIGDYVWEDVDK 540
DB 585 TTIOQTVMOTTINEYTGERTASYDNTIAFSTSSGOGGDLPPKTYKIGDYVWEDVDK 644
QY 541 DGIQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEDEKGYQFDG 582
DB 645 DGIQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEDEKGYQFDG 686

RESULT 7
ABM79020
ID ABM79020 standard; protein; 549 AA.
XX
AC ABM79020;
XX
DT 15-JAN-2004 (first entry)
XX
DE Staphylococcus epidermidis polypeptide.
XX
KW Infection; antibacterial; vaccine.
XX
OS Staphylococcus epidermidis.
XX
PN WO2003076470-A1.
XX
PD 18-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-US006415.
XX
PR 05-MAR-2002; 2002US-0361324P.
XX
PA (INH1-) INHIBITEX INC.
XX
PA (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
PI Robbins J, Vernachio J, Bowden MG;
XX
WPI; 2003-722324/68.
XX
New antibody recognizing a Staphylococcus epidermidis protein comprising
PT SdrG N1N2N3, SdrG N2N3 or SdrGR2 useful for preparing a composition for
PT treating or preventing a coagulase-negative staphylococcal infection.
XX
PS Claim 20; Page 37; 78pp; English.
XX
The present sequence comprises the protein sequence of a polypeptide of a
CC coagulase-negative Staphylococcus epidermidis. A claimed monoclonal
CC antibody recognises this protein and is used in a claimed method of
CC treating or preventing a coagulase-negative staphylococcal infection in a
CC human or animal, e.g. a nosocomial coagulase-negative staphylococcal
CC infection in low birth weight infants
XX
SQ Sequence 549 AA;

Query Match 88.6%; Score 2670; DB 7; Length 549;
Best Local Similarity 98.7%; Pred. NO. 5.7e-138;
Matches 519; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDEKNDVINNNOSINTDDNNQIIKKEETNNYDGIKRSDEDRTESTNVNDEATFLOK 60
DB 24 SSDEENDVINNNOSINDDDNNQINKKEETNNNDGIEKSSDRTESTNVNDEATFLOK 83
QY 61 TPQDNTHLTREEVEKSSVESNSSIDTAQOPSHTTINREESVOTSDNVEDSHVSDFANS 120
DB 84 SPQDNTHLTREEVEKSPSSVESNSSIDTAQOPSHTTINREESVOTSDNVEDSHVSDFANS 143
QY 121 KIKSNTESGKEENTIEOPNKKVEDSTTSQPSGYTNIDEKISNDELNLNLPINEYENKAR 180
DB 144 KIKSNTESGKEENTIEOPNKKVEDSTTSQPSGYTNIDEKISNDELNLNLPINEYENKAR 203

```

```

QY 181 PLSTTSAQPSIKRVTNQLAAEOGSNNVHLLIKVTDQSTTEGYDDSEGVIIKAHAENLIYD 240
DB 204 PLSTTSAQPSIKRVTNQLAAEOGSNNVHLLIKVTDQSTTEGYDDSEGVIIKAHAENLIYD 263
QY 241 VTFEYDDKVKSGDWTVDIDKNTVPSDLTDSFTTIPKIKDMSGELIATGTVDNKNKQIYTT 300
DB 264 VTFEYDDKVKSGDWTVDIDKNTVPSDLTDSFTTIPKIKDMSGELIATGTVDNKNKQIYTT 323
QY 301 FTDYVDKYENIKAHLLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 360
DB 324 FTDYVDKYENIKAHLLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 383
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLYSAKETNNVINSNGDEGSTIIDSTIIKVKYKVG 420
DB 384 ANLQSMFTNIDTKNHTVEQTIYINPLYSAKETNNVINSNGDEGSTIIDSTIIKVKYKVG 443
QY 421 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480
DB 444 DNQNLPSNRNIYDYSEYEDVTNDYALQGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 503
QY 481 TTIOQTVMOTTINEYTGERTASYDNTIAFSTSSGOGGDLPPK 526
DB 504 TTIOQTVMOTTINEYTGERTASYDNTIAFSTSSGOGGDLPPK 549

RESULT 8
ABM79015
ID ABM79015 standard; protein; 560 AA.
XX
AC ABM79015;
XX
DT 15-JAN-2004 (first entry)
XX
DE Staphylococcus epidermidis SdrG N1N2N3 domain.
XX
KW SdrG; surface protein; infection; antibacterial; vaccine.
XX
OS Staphylococcus epidermidis.
XX
PN WO2003076470-A1.
XX
PD 18-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-US006415.
XX
PR 05-MAR-2002; 2002US-0361324P.
XX
PA (INH1-) INHIBITEX INC.
XX
PA (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
PI Robbins J, Vernachio J, Bowden MG;
XX
WPI; 2003-722324/68.
XX
N-PSDB; ACF80624.
XX
New antibody recognizing a Staphylococcus epidermidis protein comprising
PT SdrG N1N2N3, SdrG N2N3 or SdrGR2 useful for preparing a composition for
PT treating or preventing a coagulase-negative Staphylococcal infection.
XX
PS Claim 27; Page 24-25; 78pp; English.
XX
The present sequence comprises the protein sequence of the N1N2N3 region
CC (amino acids 50-597), or putative A domain, of the SdrG surface protein
CC of coagulase-negative Staphylococcus epidermidis. A claimed antibody
CC recognises a protein selected from SdrG N1N2N3, N2N3 and TR2. The
CC antibody may be a monoclonal antibody, including a chimeric, murine,
CC humanized, human or single chain monoclonal antibody, which prevents a
CC coagulase-negative staphylococcal infection in a human or animal by
CC inhibiting binding of staphylococcus bacteria to fibrinogen. Such
CC antibodies can be used to treat or prevent staphylococcal infections
CC including nosocomial coagulase-negative staphylococcal infections in low

```


Db 124 TETKPN--IEKENVKPSDKTATDTSVILEEKKAPNNT--NNDVTTKPS-----TS 172

Qy 116 DFANSKIKESNT---ESGKEENTIEQPNKVKEDSTTSQPSGYTNI--DEKIS-NOBELNLNP 171

Db 173 EPSTSEIQTKPTTPOESTNIENSQPTPSKVD---NQVTDATPKPEPNVNSKEELKNP 229

Qy 172 -----INEYENKARPLSTTSQAQSIKEV-----TUNOLAAEQGSNNVHLIKWTDQ 216

Db 230 EKLKELVRNDSNTDHTSKPVATPTSVAPKRVNAKMFAPAQAASNNVNDLIKVTQK 289

Qy 217 SITEGYDDSEGVIKAHDAENLIYDVTFEVDKVKSGDGTMTVDIDKNTVPSDLTDSFTIPK 276

Db 290 TIKVG-DGKDNVAAAHGDKDIEYDTEFTIDNKVKGGDTMTYDKNVPSDLTDKNPID 348

Qy 277 IKDNGSEIIATGYDNNKKNQITFTYDVKYENIKAHKLKLTSDYDYSKVPNNNTKLDV 336

Db 349 ITDPSGEVIAKGTPTKATQITFTYDVKYEDIKSLRLTLYSIDKKTVP-NETSLNT 407

Qy 337 YKTALSSVNKTITVEYQRPNEANTANLOSMTNIDTKNHTVEQTIYINPLYSAKETNVN 396

Db 408 PATAGKTSQNVTVDYQDPWVHGDSNIQSIPTKLEDEKQTIHQQLYVNLPLKKSATNTKVD 467

Qy 397 ISGNG-----DEGSTIIDSTIIKVKYKVGDNQNLPSNRIDYSEVEDVTND-DYA 446

Db 468 IAGSQVDYGNIKLNGSTIIDQTEIKVKYKVSQQLPQSNRIYDSQYEDVTISQFDNK 527

Qy 447 QLGNNNDVNFNGNIDSPYIIKVKISKYPNKKDVTITQVTVMTTINEYTGEPRTASYD 506

Db 528 KSPFNNAVTLDFGINSAYIIKVKSKVTPTSDGELDIAQGTSMRTT-DKY-GYNYAGYS 585

Qy 507 NTIAPSTSSGOGGLD-PPKTYKIGDYWEDVDKDGIONTNDNEKELSNVLTLYTPDG 565

Db 586 NFVITSNDTGGDGTVPPEELKXIGDYWEDVDKDGQVGGTDSKEKPMANVLTLTPDG 645

Qy 566 TSKSVRTDECKYQFDG 582

Db 646 TTKSVRTDANGHYEFGG 662

RESULT 12

ABJ18982

ID ABJ18982 standard; protein; 1166 AA.

XX

AC ABJ18982;

XX

XX

DT 06-MAR-2003 (first entry)

XX

DE Pathogen specific antigen related staphylococcal protein SEQ ID No 153.

XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;

KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;

KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;

XX autoimmune disease; HIV; hepatitis.

OS Staphylococcus sp.

XX

XX WO200259148-A2.

PN

XX

PD 01-AUG-2002.

XX

PF 21-JAN-2002; 2002WO-EP0000546.

XX

PR 26-JAN-2001; 2001AT-00000130.

XX

XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

PA

XX Meinke A, Nagy E, Von Ahean U, Klade C, Henics T, Zauner W;

PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;

PI Tempelmaier B;

XX

XX WPI; 2003-075410/07.

XX

PT Identifying, isolating and producing hyperimmune serum-reactive antigens

PT from a pathogen, for preparing vaccine or medicament for treating or preventing e.g. staphylococcal infections, comprises providing antibody preparation.

PT

XX Example 7; Page 173; 252pp; English.

XX

CC The invention relates to a novel method for identifying, isolating and producing hyperimmune serum-reactive antigens from a pathogen, tumour, allergen, a tissue or host prone to auto-immunity, where the antigens are used in a vaccine, comprises providing antibody preparation from a plasma pool of a type of animal, or individual sera with antibodies against the specific pathogen, tumour, allergen, tissue or host prone to auto-immunity. The hyperimmune serum-reactive antigens comprising any of the 62 sequences of 53-2261 amino acids fully defined in the specification, or their hyperimmune fragments are useful for the manufacture of a pharmaceutical preparation, particularly a vaccine against staphylococcal infections or colonisation against *S. aureus* or *S. epidermidis*. The preparation of antibodies is useful for the manufacture of a medicament for treating or preventing staphylococcal infections or colonisation against *S. aureus* or *S. epidermidis*. The antibody preparations may also be used for diagnostic and imaging purposes. Other conditions that can be treated include cancer, autoimmune diseases or infections caused by viral (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This sequence represents a staphylococcal protein relating to the method for identifying and producing pathogen specific antigens of the invention

CC

XX Sequence 1166 AA;

SQ

Query Match 37.9%; Score 1142.5; DB 6; Length 1166;

Best Local Similarity 42.7%; Pred. No. 4.8e-54;

Matches 262; Conservative 107; Mismatches 196; Indels 49; Gaps 18;

Qy 1 SDEKNDVNNQSIINTDDNQI--IKKEETNNYDGIKRSSEDRSTSTNVNDNEATFL 58

Db 66 ATTSNKEVSETENNSTNNSTNPIKKE--TWDSQPEAKKESTSSSQKQNNVTAT 123

Qy 59 QKTPQDNTLTEREVKES-----SSVESNSSIDTAQPSHTTINREESVQTSNDVEDSHVS 115

Db 124 TETKPN--IEKENVKPSDKTATDTSVILEEKKAPNNT--NNDVTTKPS-----TS 172

Qy 116 DFANSKIKESNT---ESGKEENTIEQPNKVKED---STTSQPSGYTNI--DEKISNOBEL 167

Db 173 EPSTSEIQTKPTTPOESTNIENSQPTPSKVDNQVTDATNPKEPNVNSKEELKNPEKL 232

Qy 168 LNLPISE--YENKARPLSTTSQAQSIKRV-----TVNOLAAEQGSNNVHLIKWTDQSI 219

Db 233 KELVRNDSNTDHTSKPVATPTSVAPKRVNAKMFAPAQAASNNVNDLIKVTQKTIK 292

Qy 220 EGYDDSEGVIKAHDAENLIYDVTFEVDKVKSGDGTMTVDIDKNTVPSDLTDSFTIPKID 279

Db 293 VGGKDNVAAAHGDKDIEYDTEFTIDNKVKGGDTMTYDKNVPSDLTDKNPIDIDT 351

Qy 280 NSGEIIATGYDNNKKNQITFTYDVKYENIKAHKLKLTSDYDYSKVPNNNTKLDVEYKT 339

Db 352 PSGEVIAKGTPTKATQITFTYDVKYEDIKSLRLTLYSIDKKTVP-NETSLNLFAT 410

Qy 340 ALSSVNKTITVEYQRPNEANTANLOSMTNIDTKNHTVEQTIYINPLYSAKETNVNISG 399

Db 411 AGKETSQNVTVDYQDPWVHGDSNIQSIPTKLEDEKQTIHQQLYVNLPLKKSATNTKVDIAG 470

Qy 400 NG-----DEGSTIIDSTIIKVKYKVGDNQNLPSNRIDYSEVEDVTND-DYAQLG 449

Db 471 SQVDDYGNIKLNGSTIIDQTEIKVKYKVSQQLPQSNRIYDSQYEDVTISQFDNKKSF 530

Qy 450 NNDVNFNGNIDSPYIIKVKISKYPNKKDVTITQVTVMTTINEYTGEPRTASYDNTI 509

Db 531 SNNVATLDFGINSAYIIKVKSKVTPTSDGELDIAQGTSMRTT-DKY-GYNYAGYSNFI 588

Qy 510 AFSTSSGOGGLD-PPKTYKIGDYWEDVDKDGIONTNDNEKELSNVLTLYTPDGTSK 568

Db 589 VTSNDTGGDGTVPPEELKXIGDYWEDVDKDGQVGGTDSKEKPMANVLTLTPDGTTK 648

Qy 569 SVRTDECKYQFDG 582

Db 649 SVRTDANGHYEFGG 662

||||| : | : | : |

RESULT 13

ABU42327

ID ABU42327 standard; protein; 1141 AA.

XX AC ABU42327;

XX 19-JUN-2003 (first entry)

XX DT Protein encoded by Prokaryotic essential gene #27854.

XX DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX KW Staphylococcus aureus.

XX OS WO200277183-A2.

XX PN 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX N-PSDB; ACA46197.

DR WPI; 2003-029926/02.

XX New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 70251; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX CC Sequence 1141 AA;

XX Query Match 37.8%; Score 1137.5; DB 6; Length 1141;

XX Best Local Similarity 43.3%; Pred. No. 8.8e-54;

XX Matches 266; Conservative 97; Mismatches 193; Indels 59; Gaps 19;

Qy 1 SSDEEKNDVINNQSINTDDNNQIIKKEETNNVDGIEKRSDETESTNVDENATFLOK 60

Db 69 SDNKEVSVETENNSTTENDSTNPI--KKEINT--DSQPEAKEESTTSQQQNNVTATTE 125

Qy 61 TPQDNTHLTREEVKES---SSVESNSSIDTAQOPSHTTINREESVCTSDNVEDSHVSDF 117

Db 126 TKQN--IEKNVKPSTDKTATEDTSVILEKAPNT--NNDVTYKPS----- 170

Qy 118 ANSKIKESNT---ESGKEENTIEQPNKVEDSTTSQSGYTNIDEKIS-NQDELLNLP-- 171

Db 171 -TSEIQTKPTTPOESTNIENSQPQTPSKVD--NQVTDATNPKEPVNVSKELKNPEK 226

Qy 172 -----INEYENKARPLST--TSAQP-----SIKRVTVNQAAEQGSNVNHLIKVTDQSI 218

Db 227 LKELVRNDNNTDRSTKPVATAPTSAVKRLNAXKRFAPAQAASNNVNDLITVTKQTI 286

Qy 219 TEGYDDSEGVKAHDAENLIYDVTFEVDKVGSDTMTVDIDKNTVPSDLTDSFTIPKIK 278

Db 287 KVG-DGKDNVAAADGKDIETDTEFTIDNKVKGDFTWINDKNTVPSDLTDRKNDPIDIT 345

Qy 279 DNSGEIATGTYNKNKQIITYTFTDVKYKENTKAHLKLTYSIDKSKVPNNNTKLDVYK 338

Db 346 DPSGEVIAKGTDFKATKQIITYTFTDVKYEDIKARLTLYSIDKQAVP-NETSLNLTEA 404

Qy 339 TALSSVNNKTTTVEYQRPENERTANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNNNIS 398

Db 405 TAGKETSONVSDYQDPMVHGDSNIQSIFFKLDEKQTEEQIYVNPVKKTATNTKVDIA 464

Qy 399 GNG-----DEGSIIDDSIIKVKYKVDGNQNLPSNRIYDYSEYEDVTND-DYAOQL 448

Db 465 GSQVDDYGNIKLNGSTIIDQNTETIKVKYKVPNQQLPQSNRIYDFSQYEDVTSQDFNKS 524

Qy 449 GNNNDVINFGNIDSPYIIKVKSKYDNPKNDDYTTIOQVTMTQTTINEYGEFTASYDNT 508

Db 525 FSNVATLDFGNDINSAYIIKXWSKYTPTSDGELDIAQGTSMRTT-DKY-GYNYAGYSNF 592

Qy 509 IAFSTSSGQOQDL-PPEKTYKIGDYVWEDVDKQNTNDNEKPLSNVLVTITYPDGTS 567

Db 583 IVTSNDTGGDGTVPKEEKLKIGDYVWEDVDKQGTGDSKEKPMANVLITYPDGTT 642

Qy 568 KSVRTDEDEGKYQPDG 582

Db 643 KSVRTDANGHYEFGG 657

RESULT 14

ABU42520

ID ABU42520 standard; protein; 670 AA.

XX AC ABU42520;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #28047.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Staphylococcus epidermidis.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX XX
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA46390.
XX XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX XX
XX PS Claim 25; SEQ ID NO 70444; 1766pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: the sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 670 AA;
Query Match 37.1%; Score 1116.5; DB 6; Length 670;
Best Local Similarity 98.2%; Pred. No. 6.1e-53;
Matches 213; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
Qy 366 MFTNIDKNTVEQTIYNPLRYSAKETNNVNSGNDEGSGTIIDSTIIKVKYKVDNQNL 425
Db 1 MFTNIDKNTVEQTIYNPLRYSAKETNNVNSGNDEGSGTIIDSTIIKVKYKVDNQNL 60
Qy 426 PDSNRIYDSEYEDVNDVYDVAQLGNNNDVNFNGNIDSPYIIKVIKYDPNKDDYTTIIQ 485
Db 61 PDSNRIYDSEYEDVNDVYDVAQLGNNNDVNFNGNIDSPYIIKVIKYDPNKDDYTTIIQ 120
Qy 486 TVTWTQTINETYGFRFASVDNTIAFTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 545
Db 121 TVTWTQTINETYGFRFASVDNTIAFTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 177
Qy 546 TNDNEKPLSNVLVTLPDGTSKSVRTDEDEKGYQFDG 582
|||||

Db 178 TNDNEKPLSNVLVTLPDGTSKSVRTDEDEKGYQFDG 214
RESULT 15
AAG82803
ID AAG82803 standard; protein; 278 AA.
XX AC AAG82803;
XX DT 03-SEP-2001 (first entry)
XX XX
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:2700.
XX XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
XX KW endocarditis.
XX XX Staphylococcus epidermidis.
XX PN WO200134809-A2.
XX PD 17-MAY-2001.
XX PF 09-NOV-2000; 2000WO-US030782.
XX PR 09-NOV-1999; 99US-0164258P.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Kimmerly WJ;
XX XX WPI; 2001-316495/33.
XX DR N-PSDB; AAH53653.
XX XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX XX useful for vaccinating against infections, e.g. endocarditis.
XX PS Claim 18; Page 705; 2188pp; English.
XX XX
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
XX CC and (II) can have antibacterial activity and therefore can be used in
XX CC vaccination. The nucleic acids (I) may be used to produce the S.
XX CC epidermidis polypeptides (II) via the production of vectors containing
XX CC them which are used to produce hosts cells which express the
XX CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX CC used to vaccinate subjects and to raise antibodies against the bacteria.
XX CC The polypeptides may also be used to assay for other inhibitors of their
XX CC activity and therefore identify compounds that may be used for the
XX CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
XX CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
XX CC represent oligonucleotide sequences and primers which are used in the
XX CC exemplification of the present invention. N.B. The present invention
XX CC specifically claims all the polynucleotide sequences given in the
XX CC sequence listing of the present specification, however the sequence
XX CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
XX CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
XX CC for SEQ ID NO:4455 to 4464
XX SQ Sequence 278 AA;
Query Match 28.4%; Score 857; DB 4; Length 278;
Best Local Similarity 93.0%; Pred. No. 2.9e-39;
Matches 172; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
Qy 1 SSDEKNDVNNQSNITDNNNOIIEKEETNNYDGIKRSSEDRTESTTNVDENATFLQK 60
Db 75 SSDEENDVNNQSNITDNNNOIIEKEETNNYDGIKRSSEDRTESTTNVDENATFLQK 134
Qy 61 TPQDNTHLTETEEVKESSESSVESSNSSIDTAQPSHTTINREESVQTSNVDSDSHVDFANS 120
Db 135 SPQDNTHLTETEEVKESSESSVESSNSSIDTAQPSHTTINREESVQTSNVDSDSHVDFANS 194
|||||

Qy 121 KIKESNTESGKEENTIEQENKVKEDSTTSQPSGYTNIDEKISNODELLNLPINEYENKAR 180
Db 195 KIKESNTESGKEENTIEQENKVKEDSTTSQPSGYTNIDEKISNODELLNLPINEY--KVT 252
Qy 181 PLSTT 185
Db 253 KLSVT 257

Search completed: October 5, 2004, 20:48:05
Job time : 70.362 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 17.0358 Seconds
(without alignments)
1763.712 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEEKNDVINNQSIINTDD.....PDGTSKSVRTDECKYQFDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2810.5	93.3	930	4	US-09-134-001C-5314
2	2810.5	93.3	930	4	US-09-386-962C-10
3	1144.5	38.0	1166	4	US-09-200-650E-7
4	576	19.1	1315	4	US-09-200-650E-5
5	499	16.6	1742	4	US-09-386-962C-4
6	455	15.1	930	4	US-09-200-650E-3
7	454	15.1	936	4	US-08-956-171E-5249
8	452	15.0	933	3	US-08-293-728-2
9	452	15.0	933	3	US-09-421-868-2
10	396	13.1	918	4	US-09-200-650E-1
11	349	11.6	345	3	US-08-856-253-7
12	300.5	10.0	1027	4	US-08-956-171E-5254
13	259.5	8.6	1112	2	US-08-714-402-2
14	259.5	8.6	1161	4	US-09-327-536-2
15	220.5	7.3	3696	4	US-09-134-001C-5080
16	209	6.9	630	3	US-08-973-462-9
17	208.5	6.9	251	4	US-08-956-171E-5252
18	203.5	6.8	2504	4	US-09-328-352-5821
19	202.5	6.7	368	4	US-09-134-000C-4597
20	202	6.7	676	4	US-09-134-001C-4318
21	201.5	6.7	1177	4	US-09-134-001C-5106
22	197	6.5	2314	4	US-09-268-347-49
23	195	6.5	1335	4	US-09-134-001C-3716
24	195	6.5	10182	4	US-09-134-001C-3159
25	192.5	6.4	1306	4	US-09-134-000C-6670
26	192.5	6.4	1786	3	US-08-973-462-8
27	190.5	6.3	1833	4	US-08-621-944A-4

28	190.5	6.3	1833	4	US-08-945-567D-4	Sequence 4, Appli
29	190.5	6.3	1992	4	US-08-621-944A-3	Sequence 3, Appli
30	190.5	6.3	1992	4	US-08-945-567D-3	Sequence 3, Appli
31	189.5	6.3	1664	1	US-09-599-652-2	Sequence 2, Appli
32	189.5	6.3	1664	2	US-08-642-846-2	Sequence 2, Appli
33	189.5	6.3	1664	4	US-09-264-604-2	Sequence 2, Appli
34	188.5	6.3	1435	2	US-08-568-459A-4	Sequence 4, Appli
35	188.5	6.3	1435	2	US-08-487-826B-4	Sequence 4, Appli
36	188.5	6.3	1435	4	US-09-210-288-4	Sequence 4, Appli
37	187	6.2	1416	4	US-09-071-035-404	Sequence 404, App
38	187	6.2	1448	4	US-09-071-035-402	Sequence 402, App
39	186.5	6.2	3088	4	US-09-562-702A-8	Sequence 8, Appli
40	186.5	6.2	3089	4	US-09-562-702A-4	Sequence 4, Appli
41	186.5	6.2	3110	4	US-09-562-702A-2	Sequence 2, Appli
42	186.5	6.2	3110	4	US-09-562-702A-6	Sequence 6, Appli
43	186.5	6.2	3110	4	US-09-561-709B-7	Sequence 7, Appli
44	186.5	6.2	3111	2	US-08-460-309-4	Sequence 4, Appli
45	186.5	6.2	3111	2	US-08-125-077-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match	93.3%	Score	2810.5	DB	4	Length	930
Best Local Similarity	93.6%	Pred. No.	2e-158				
Matches	545	Conservative	15	Mismatches	21	Indels	1
Gaps	1						
Qy	1	SSDEEKNDVINNQSIINTDDNQIIKKEETNNVDGIEKRSIEDRTSTNTVDENEATFLOK	60				
Db	73	SSNEEKNDVINNQSIINTDDNQ-IKKEETNSNDIAENRSKDIQTSTNTVDENEATFLOK	131				
Qy	61	TPQDNTHLTEEVKESSESSVSSSIDTAOPSHHTINREESVOTSDNVEDSHVSDPANS	120				
Db	132	TPQDNTQLKEEVKPESSVSSSSMDTAOPSHHTINSEASIQTSNEENSRVSDPANS	191				
Qy	121	KIKESNTESKCENTIEQPNKVKEDSTTSQPSGYTNIDEXISNQDELNLNLPINEYENKAR	180				
Db	192	KIESNTESKCENTIEQPNKVEDSITSQPSKYKIDEXISNQDELNLNLPINEYENKVR	251				
Qy	181	PLSTTSAPQSIKRVTVNQLAAEQSGSNVHLIKVTDQSITEGYDDSEGVKAHAENLIYD	240				
Db	252	PLSTTSAPQSKRVTVNQLAAEQSGSNVHLIKVTDQSITEGYDDSDGIIKAHAENLIYD	311				
Qy	241	VTFEVDKVKSGDVTVDIDKNTVPDLTDSFTIPKIKDNGSETIATGTVDNKNQIITYT	300				
Db	312	VTFEVDKVKSGDVTVDIDKNTVPDLTDSFATPKIKDNGSETIATGTVDNKNQIITYT	371				
Qy	301	FTDVKYKNIKAHLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPENRT	360				
Db	372	FTDVKYKNIKAHLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPENRT	431				

QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGDEGSTIIDSTIIKVKYVG 420
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGDEGSTIIDSTIIKVKYVG 491
QY 421 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNDVNIINFNGIDSPYIIKVIKIDYDPNKDYY 480
Db 492 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNDVNIINFNGIDSPYIIKVIKIDYDPNKDYY 551
QY 481 TTIQQTVMQTTINEYTGFEPTASDYNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDK 540
Db 552 TTIQQTVMQTTINEYTGFEPTASDYNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDK 611
QY 541 DGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582
Db 612 DGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 653

RESULT 2

US-09-386-962C-10
; Sequence 10, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-962C-10

Query Match 93.3%; Score 2810.5; DB 4; Length 930;
Best Local Similarity 93.6%; Pred. No. 2e-158;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;
QY 1 SSDEKNDVINNQSDINTDNNQIIKKEETNNYDIEKRSBDRTESTTNVDENAEATFLQK 60
Db 73 SSNEEKNDVINNQSDINTDNNQIIKKEETNSDALENRSKDIQTSTNVDENEAEATFLQK 131
QY 61 TPQDNTHLTETBVKSSSVESNSSIDTAQOPSHHTTNREESVQTSNVEDSHVSDFANS 120
Db 132 TPQDNTHLTETBVKSSSVESNSSMDTAQOPSHHTTNREESVQTSNVEDSHVSDFANS 191
QY 121 KIKESNTESGKEENTIEQPNVKEDSTTSQPSGYTNIDEKISNQDELINLPINEYENKAR 180
Db 192 KIKESNTESGKEENTIEQPNVKEDSTTSQPSGYTNIDEKISNQDELINLPINEYENKAR 251
QY 181 PLSTTSQAQPSKRVTVNQLAAEQSGNVNHLIKVTDQSTIEGYDSEGVKAHAENLIYD 240
Db 252 PLSTTSQAQPSKRVTVNQLAAEQSGNVNHLIKVTDQSTIEGYDSEGVKAHAENLIYD 311
QY 241 VTFEVDDKVKSGDVTVDIDKNTVPSDLTSDFTIPKIKNSGELIATGVNKNKOITYT 300
Db 312 VTFEVDDKVKSGDVTVDIDKNTVPSDLTSDFTIPKIKNSGELIATGVNKNKOITYT 371
QY 301 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKNTALSSVKNKTIIVEYQRPENRT 360
Db 372 FTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVEYKNTALSSVKNKTIIVEYQRPENRT 431
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGDEGSTIIDSTIIKVKYVG 420
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGDEGSTIIDSTIIKVKYVG 491
QY 421 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNDVNIINFNGIDSPYIIKVIKIDYDPNKDYY 480

Db 492 DNQNLPSNRIYDYSEYEDVTNDYDIAQLGNNDVNIINFNGIDSPYIIKVIKIDYDPNKDYY 551
QY 481 TTIQQTVMQTTINEYTGFEPTASDYNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDK 540
Db 552 TTIQQTVMQTTINEYTGFEPTASDYNTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDK 611
QY 541 DGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582
Db 612 DGIONTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 653

RESULT 3

US-09-200-650E-7
; Sequence 7, Application US/09200650E
; Patent No. 6680195
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eichmann, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-7

Query Match 38.0%; Score 1144.5; DB 4; Length 1166;
Best Local Similarity 42.6%; Pred. No. 7.9e-60;
Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;
QY 1 SSDEKNDVINNQSDINTDNNQIIKKEETNNYDIEKRSBDRTESTTNVDENAEATFL 58
Db 66 ATTSNKNVSEVSESTENNSTNNSTNPIKKE--TNTDSQPEAKKESSTSSQKQNNVTAT 123
QY 59 QKTPQDNTHLTETBVKES---SSVESNSSIDTAQOPSHHTTNREESVQTSNVEDSHVS 115
Db 124 TETKPN--IEKENVPSDTKATEDTSVILEKAPNNT--NNDVTTKPS-----TS 172
QY 116 DFANSKIKESNT---ESGKEENTIEQPNVKEDSTTSQPSGYTNIDEKIS-NODELINLP 171
Db 173 EPSTSEIQTKPTTPQSTNLIENSQOPTPSKVD---NQVTDATNPKEPVNVSKELKNP 229
QY 172 -----INVENKARPLSTTSQAQPSKRV-----TVNQLAAEQSGNVNHLIKVTDQ 216
Db 230 EKLKELVRNDSNTDHTSKPVATAPTSVAPKRVNAKMFAPAAPAAVANNVNDLIKVTQ 289
QY 217 SITEGVDDSEGVKAHAENLIYDVTFEVDKVKSGDVTVDIDKNTVPSDLTSDFTIPK 276
Db 290 TIKVG-DGKNVAAAHGDKDIEYDTEFTIDNKNVKGDTMTINFKNVIIPSDLTCKNDPID 348
QY 277 IKDNGEIIATGYDNNKQIYTFYTDYVDKYENIKAHKLKTSYIDSKVPNNNTKLDVE 336
Db 349 ITDPSGEVIAKGFDFDKATQIYTFYTDYVDKYEDIKSRLTLYSIDKTVP-NETSLNT 407
QY 337 YKTALSSVKNKTIIVEYQRPENRTALOSMFTNIDTKNHTVEQTIYINPLRYSAKETNWN 396
Db 408 FATAGKETSNVTVDYQDPMVHGDSNIQSIIFTKLDEKQTIIEQIYVNFPLKKSATNTKVD 467
QY 397 ISGNG-----DRGSTIIDSTIIKVKYGVNDNQLPSNRIYDYSEYEDVTND-DYA 446
Db 468 IAGSVDDYGNIKLNGKSGSTIIIDQNTIEKVKYKNSDQQLPQSNRIYDFSQYEDVTSQFDNK 527

Db 84 TAKUPPIWAGD-QVLANGVIDS-DGNVIYTFDVTYVNTKDDVKATLTMPAYID-----PENV 137
Qy 331 TKL-DVEYKTALSS--VNKTIIVYQRPENRNTANLQSMFTNIDTKNHTVEQTTIYNPL- 386
Db 138 KKTGNVLTATGIGSTANKTANLVLDYKYGKFPYNSLKGITDIDKTNNTYRQTTIYNPSG 197
Qy 387 -RYSAKETNVMISNGDEGSIIDDSITIKYKVGDNQNLPSNRIVYDSEYEDVTDDY 445
Db 198 DNVIAPVLTKNLKPTNSNALIDQNTSIKYKYVDNAADLSESYFV-NPENFEDVTNSVN 256
Qy 446 AQLGNNDVNINFG---NIDSPYIKVSKYDPN-KDDYTTIQOTVTMOTTINEXYTG- 499
Db 257 ITFNPQYKVFNTPDQITPTPIVVVNGHIDPNSKGD-----LAIRSTLYGYSNI 309
Qy 500 -FRTASYDNTIAFTSSGQGG-DLP--PEK 526
Db 310 IWRSMWDEVAFNNGSGGIDKPVVPEQ 340

RESULT 12

US-08-956-171E-5254
; Sequence 5254, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Pannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hymen
; REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5254:
; SEQUENCE CHARACTERISTICS:

LENGTH: 1027 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5254:

US-08-956-171E-5254

Query Match 10.0%; Score 300.5; DB 4; Length 1027;
Best Local Similarity 21.2%; Fred.No. 5.1e-10;
Matches 123; Conservative 118; Mismatches 243; Indels 97; Gaps 25;

Qy 35 GIEKRSBDRTESTNTVDENEAFLQKTPQN-----THLTVEEVKESSESVSSSSSIDTAQ 90
Db 39 GQKKEAAASEQKTTTVEEN-----GNSATDNKTSSETOTTATNVNHIETQSYNATV--TE 91
Qy 91 QPSHTT-INREER---SVQSDNVEDSHVDFANSKIKESNTESGKEENTLEQPNKVKEDS 146
Db 92 QPSNATQVTTTEAPKVAQAPQATQAPANIETVKEBWK-----EAKPQVKETT 139
Qy 147 TTSOPSG---YTNIDKISNQDELLNLPINEYENKARPLSTSAQPSIKRVT-----V 196
Db 140 QSDQNSGQDQVLDLTPKATQNVAEQV---EVAQPTASESKPRVTRSDVAEAKA 195
Qy 197 NQLAAREGSGNVNHLIKVDQSITEGYDDSDSGVTKAHDENLIYDVTVEDDKVKSGDTMT 256
Db 196 SNAKVETGTDVTSKVTVEIGSI-EGHNNTNKV-EPHAGORAVLYKLFKFNGLHQGDYFD 253
Qy 257 VDIDKNTVPSDLTDSFTPIPKDNS-----GEIIATGYDNKNKIITYTPTDYVDKYENI 311
Db 254 FTLSNNVNTGHVSTARKVPEIKNGSVNATGEVLGG-----KIRYTFNDIEDKVDV 306
Qy 312 KAHLKLTSYIDKSKVPNNNTKLDVEYKLTALSSVN-----KTIITVEYQRPENRNTANLQSM 366
Db 307 TAELEINLFDPKTVQTNNGN-----QTITSTLNEEQTSKELDVKKYKDGIGNYANLGS 360
Qy 367 FTNIDTKNHTVEQTIYINPLRYSAKETNVMISNGDEGSIIDDSITIKVKYK-VGDQNVL 425
Db 361 IETFNKANRRFHVAFIKP--NNGKTTSVTVTGLMKSGNQNGQPKVRIPEYLGNNEDI 418
Qy 426 PDSNRIV---DYSEVEDVTNDYAOGL--NNNDVNINFGNIDSPYIIKVISKYDPNKO- 478
Db 419 AKS--VYANTTDTSKFKEVTSNMGSLNQLNGSGYSNIENLDKTYVYVHDGEYLVGTDE 476
Qy 479 -DYTTIQOTVTMOTTINNEYTGEFTASYDNTIAFTSSGQGGDLPP-----EKTYKIGD 532
Db 477 VDFRT-QMVGHPEQLYKYYYDRGYTLTWDLNGLVLSKANGNEKNGPIIQNNKFEYK--- 532
Qy 533 YVMEDVDKGIQNTNDNEKPLSNVLTLTVPDGTSKSVRTD 573
Db 533 ---EDTIKETLTGQYDKN-----LVTTVBEEYDSDSTLDD 564

RESULT 13

US-08-714-402-2
; Sequence 2, Application US/08714402
; Patent No. 5910441
; GENERAL INFORMATION:
; APPLICANT: ROCHA, Claudia
; APPLICANT: FISCHETTI, Vincent A.
; TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING
; TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,402
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 016921-097
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-714-402-2

Query Match 8.6%; Score 259.5; DB 2; Length 1112;
Best Local Similarity 21.2%; Pred. No. 1.5e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;
QY 125 SNTSGKEENTIEQPNKVKEDSTTSQPS-----GYTNIDEKISNQDELL--NLPIN 173
DB 43 ANKQGSFEIKKVDQNNKPLPGATSSLTSDKGKTSVQSFTSNDKGIYDAQNLPQGTTLK 102
QY 174 E-----YENKARPLSTTSQAQPSIKRVTVN-----QLAAEQGS-NVNHLIKVTD--QSITE 220
DB 103 EETAPDGYDKTSRTWTVTYVYENGTYKLVENPYNGEIIISKAGSKDVSSSLQLENPKMSVVS 162
QY 221 GYDSE--GVIKAHDAENLIYDVTFEVDKVKVKS-----GDTMTVDIDKNTVPSDLTDSF 272
DB 163 KYGKTEVSSGAADFYRNHAAFYKMSFELKQDKKSETINPGDTFVLQDLRLNPKGISQD- 221
QY 273 TIPK-IKNSGHEIATGYDNKNKQITFTDYVDKYENIKAHKLKLTYSIDKSKVPPNNT 331
DB 332 KLDVEYKTALSSVKNKTI-----TVYEQRPNE-----NRTANLQSMFTNIDTK 373
DB 279 -----TSISNFKSTIGQGEITYKGTIVNVLYGNESTKESNYITNGLSNVGGIESYNTE 331
QY 374 NHTVEQTIYNPLRYSAKETNVNLSNG-----DEGSIIDDSITIIKVKVGDNQ 423
DB 332 TGEFVWYVYVNPRTNIPYATNMLGFGARSNSTDLNDANTSSAELGEIQVTEVPEGE 391
QY 424 NLPDSNRIDYSEYEDVT-----NDVYAOGLNNNDV-----NINFGN--IDSPYIIKVISK 472
DB 392 KLPSSYGV-----DVTKLTLRTDTAGLNGFGMTKQRIDFGNNIKNKAFIIKVTGK 444
QY 473 YDPNKDDYTIQQTVMQTTINEYTGERTAS-----YDNTIAFSTSSGQGG--- 520
DB 445 TDQS-----GKPLVVQSNLASFRGASEYAAFTVPGVGNVYFQNEIALSPSKSGSGKSE 497
QY 521 ----DLPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTYPDGTSKSVR--TDE 574
DB 498 FTKPSITVANLKRVAQLAFKKMSTDNV-----PLPEAAFEIIRSSNGNSQKLEASNT 549
QY 575 DGKYQF 580
DB 550 QGEVHF 555

RESULT 14
US-09-327-536-2
Sequence 2, Application US/09327536
Patent No. 6355477
GENERAL INFORMATION:
APPLICANT: FISCHETTI, Vincent A.
APPLICANT: ROCHA, Claudia
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A
FILE REFERENCE: 022927-008
CURRENT APPLICATION NUMBER: US/09/327,536
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: US 08/714,402
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2

LENGTH: 1161
TYPE: PRT
ORGANISM: SFPBP gene
US-09-327-536-2

Query Match 8.6%; Score 259.5; DB 4; Length 1161;
Best Local Similarity 21.2%; Pred. No. 1.6e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;
QY 125 SNTSGKEENTIEQPNKVKEDSTTSQPS-----GYTNIDEKISNQDELL--NLPIN 173
DB 43 ANKQGSFEIKKVDQNNKPLPGATSSLTSDKGKTSVQSFTSNDKGIYDAQNLPQGTTLK 102
QY 174 E-----YENKARPLSTTSQAQPSIKRVTVN-----QLAAEQGS-NVNHLIKVTD--QSITE 220
DB 103 EETAPDGYDKTSRTWTVTYVYENGTYKLVENPYNGEIIISKAGSKDVSSSLQLENPKMSVVS 162
QY 221 GYDSE--GVIKAHDAENLIYDVTFEVDKVKVKS-----GDTMTVDIDKNTVPSDLTDSF 272
DB 163 KYGKTEVSSGAADFYRNHAAFYKMSFELKQDKKSETINPGDTFVLQDLRLNPKGISQD- 221
QY 273 TIPK-IKNSGHEIATGYDNKNKQITFTDYVDKYENIKAHKLKLTYSIDKSKVPPNNT 331
DB 222 -IPKLIYDSANSPLAIGKHAENHQLIYFTDYIAGLQVLSAELSLFLENKEVLEN-- 278
QY 332 KLDVEYKTALSSVKNKTI-----TVYEQRPNE-----NRTANLQSMFTNIDTK 373
DB 279 -----TSISNFKSTIGQGEITYKGTIVNVLYGNESTKESNYITNGLSNVGGIESYNTE 331
QY 374 NHTVEQTIYNPLRYSAKETNVNLSNG-----DEGSIIDDSITIIKVKVGDNQ 423
DB 332 TGEFVWYVYVNPRTNIPYATNMLGFGARSNSTDLNDANTSSAELGEIQVTEVPEGE 391
QY 424 NLPDSNRIDYSEYEDVT-----NDVYAOGLNNNDV-----NINFGN--IDSPYIIKVISK 472
DB 392 KLPSSYGV-----DVTKLTLRTDTAGLNGFGMTKQRIDFGNNIKNKAFIIKVTGK 444
QY 473 YDPNKDDYTIQQTVMQTTINEYTGERTAS-----YDNTIAFSTSSGQGG--- 520
DB 445 TDQS-----GKPLVVQSNLASFRGASEYAAFTVPGVGNVYFQNEIALSPSKSGSGKSE 497
QY 521 ----DLPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVTLTYPDGTSKSVR--TDE 574
DB 498 FTKPSITVANLKRVAQLRFKKMSTDNV-----PLPEAAFEIIRSSNGNSQKLEASNT 549
QY 575 DGKYQF 580
DB 550 QGEVHF 555

RESULT 15
US-09-134-001C-5080
Sequence 5080, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5080
LENGTH: 3696
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080
Query Match 7.3%; Score 220.5; DB 4; Length 3696;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:49:18 ; Search time 68.491 Seconds
(without alignments)
2734.476 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEEKNDVNNQNSINTDD.....PDGTSKSVRTDEGKYQFDG 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCN_NEW_PUB.pap:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCUTS_PUBCOMB.pap:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	15	US-10-378-674-8
2	2810.5	93.3	892	12	US-10-282-122A-70481
3	2810.5	93.3	892	16	US-10-661-809-21
4	2810.5	93.3	930	12	US-10-615-383-10
5	2810.5	93.3	930	16	US-10-689-082-10
6	2810.5	93.3	930	15	US-10-690-184-10
7	2670	88.6	549	15	US-10-378-674-9
8	2510.5	83.3	560	15	US-10-378-674-2
9	1672	55.5	343	15	US-10-378-674-4
10	1562	51.8	316	15	US-10-378-674-6
11	1137.5	37.1	1141	12	US-10-282-122A-70251
12	1116.5	37.1	670	12	US-10-282-122A-70444
13	576	19.1	1349	9	US-09-815-242-5898
14	576	19.1	1349	9	US-09-815-242-13137
15	559	18.6	1385	12	US-10-282-122A-44324

16	499.5	16.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
17	499.5	16.6	932	9	US-09-815-242-12438	Sequence 12438, A
18	499	16.6	1742	12	US-10-615-383-4	Sequence 4, Appli
19	499	16.6	1742	16	US-10-689-082-4	Sequence 4, Appli
20	499	16.6	1742	16	US-10-690-184-4	Sequence 4, Appli
21	493	16.4	1633	12	US-10-282-122A-70437	Sequence 70437, A
22	489	16.2	953	12	US-10-282-122A-44457	Sequence 44457, A
23	486.5	16.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
24	455	15.1	841	9	US-09-815-242-5779	Sequence 5779, Ap
25	455	15.1	841	9	US-09-815-242-12751	Sequence 12751, A
26	454	15.1	936	8	US-08-781-986A-5249	Sequence 5249, Ap
27	454	15.1	936	12	US-10-329-624-5249	Sequence 5249, Ap
28	452	15.0	935	12	US-10-282-122A-44326	Sequence 44326, A
29	449	14.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
30	449	14.9	1021	9	US-09-815-242-12544	Sequence 12544, A
31	443.5	14.7	520	14	US-10-056-052-2	Sequence 2, Appli
32	397.5	13.2	877	12	US-10-282-122A-70428	Sequence 70428, A
33	377	12.5	496	14	US-10-311-879-29	Sequence 29, Appli
34	376	12.5	767	9	US-09-815-242-5899	Sequence 5899, Ap
35	376	12.5	767	9	US-09-815-242-13140	Sequence 13140, A
36	349	11.6	331	14	US-10-056-052-4	Sequence 4, Appli
37	349	11.6	345	9	US-09-813-820-7	Sequence 7, Appli
38	344.5	11.4	961	12	US-10-282-122A-43778	Sequence 43778, A
39	326	10.8	1038	12	US-10-282-122A-43827	Sequence 43827, A
40	319.5	10.6	978	9	US-09-815-242-5456	Sequence 5456, Ap
41	319.5	10.6	1001	9	US-09-815-242-12686	Sequence 12686, A
42	308.5	10.2	1018	9	US-09-815-242-5797	Sequence 5797, Ap
43	308.5	10.2	1018	9	US-09-815-242-12838	Sequence 12838, A
44	300.5	10.0	1027	8	US-08-781-986A-5254	Sequence 5254, Ap
45	300.5	10.0	1027	12	US-10-329-624-5254	Sequence 5254, Ap

ALIGNMENTS

RESULT 1

US-10-378-674-8
; Sequence 8, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PARTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; TYPE: PRT
; LENGTH: 1092
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-8

Query Match	100.0%;	Score 3013;	DB 15;	Length 1092;
Best Local Similarity	100.0%;	Pred. No. 1.3e-176;		
Matches 582;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SSDEEKNDVNNQNSINTDDNNQIIKKEETNNYDGLTEKRSDETESTTNDENATFLQK	60	
Db	75	SSDEEKNDVNNQNSINTDDNNQIIKKEETNNYDGLTEKRSDETESTTNDENATFLQK	134	
Qy	61	TPQDNTHLTEEEVKESSESSVSSSSIDTAQPSHTTINREESVQTSNVDSDSHVSDPANS	120	
Db	135	TPQDNTHLTEEEVKESSESSVSSSSIDTAQPSHTTINREESVQTSNVDSDSHVSDPANS	194	
Qy	121	KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELNLNLPINEYENKAR	180	
Db	195	KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELNLNLPINEYENKAR	254	
Qy	181	PLSTTSQPSIKRVTVNNQLAAEQGSNNVHLIKVTDQISITGEYDDSEGVIAKHAENLIYD	240	

Db 255 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITGEYDDSEGVIKAHDAENLIYD 314
Qy 241 VTFEVDDKVKSGDVTWIDKNTVPSDLTDSFTPIKIDNSGEIITATGTYNKNKQIITYT 300
Db 315 VTFEVDDKVKSGDVTWIDKNTVPSDLTDSFTPIKIDNSGEIITATGTYNKNKQIITYT 374
Qy 301 FTDYVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 360
Db 375 FTDYVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 434
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 420
Db 435 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 494
Qy 421 DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPVIIKVIKYPNKDDY 480
Db 495 DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPVIIKVIKYPNKDDY 554
Qy 481 TTIQQTVMQTTINEYTGFEFTASYDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 540
Db 555 TTIQQTVMQTTINEYTGFEFTASYDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 614
Qy 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582
Db 615 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 656

RESULT 2

US-10-282-122A-70481
; Sequence 70481, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70481

; LENGTH: 892
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-70481

Query Match 93.3%; Score 2810.5; DB 12; Length 892;
Best Local Similarity 93.6%; Pred. No. 2.8e-164;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKRSERDSTESTNVDHEATLQK 60
Db 35 SSNEKNDVNNQSIINTDDNQ-IKKEETNSDAIENRSKOITQSTNVDHEATLQK 93
Qy 61 TPQDNTLHTEEEVKESSESSSSIDTAQOPSHHTINREESVQTSNDVSDSHVDFANS 120
Db 94 TPQDNTLKEEVKESSESSSSMDTAQOPSHHTINSEASIQTSNEENSRSVDFANS 153
Qy 121 KIKESNTESKENTIEQPNKVKEDSTSPSGVTNIDEKISNQDELLNLPINEYENKAR 180
Db 154 KIIESNTESKENTIEQPNKVEDSITSQSPSYKXNIDEKISNQDELLNLPINEYENKVR 213
Qy 181 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITGEYDDSEGVIKAHDAENLIYD 240
Db 214 PLSTTSAQPSIKRVTNQLAAEQGSNVNHLIKVTDQSITGEYDDSEGVIKAHDAENLIYD 273
Qy 241 VTFEVDDKVKSGDVTWIDKNTVPSDLTDSFTPIKIDNSGEIITATGTYNKNKQIITYT 300
Db 274 VTFEVDDKVKSGDVTWIDKNTVPSDLTDSFAPKIDNSGEIITATGTYNKNKQIITYT 333
Qy 301 FTDYVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 360
Db 334 FTDYVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTALSSVNNKTITVEYQRPENRT 393
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 420
Db 394 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGDEGSIIDSTIIKVKYVG 453
Qy 421 DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPYIIKVIKYPNKDDY 480
Db 454 DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNINFGNIDSPYIIKVIKYPNKDDY 513
Qy 481 TTIQQTVMQTTINEYTGFEFTASYDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 540
Db 514 TTIQQTVMQTTINEYTGFEFTASYDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 573
Qy 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582
Db 574 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 615

RESULT 3

US-10-661-809-21
; Sequence 21, Application US/10661809
; Publication No. US20040101919A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; FILE REFERENCE: P07741US01/BAS
; CURRENT APPLICATION NUMBER: US/10/661,809
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/410303
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-661-809-21

Query Match 93.3%; Score 2810.5; DB 16; Length 892;
Best Local Similarity 93.6%; Pred. No. 2.8e-164;

Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 60
Db 35 SSNEEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 93

Qy 61 TPQDNTHLTETEEVKESSESSVSSSSIDTAQOPSHHTINREESVQTSNVDSDSHVDFPANS 120
Db 94 TPQDNTHLTETEEVKESSESSVSSSSMDTAQOPSHHTINSEASIQTSNDEENSRVDFPANS 153

Qy 121 KIKESNTESGKEENTIEQPNKVKEDSTTSOPSGYTNIDDKISNODELNLPINEYENKAR 180
Db 154 KIKESNTESGKEENTIEQPNKVRDSTTSOPSSYKNIDDKISNODELNLPINEYENKVR 213

Qy 181 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHLLIKVTDQSIETGYDDSEGVIIKAHDAENLIYD 240
Db 214 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHLLIKVTDQSIETGYDDSGI IKAHDAENLIYD 273

Qy 241 VTFEVDKVKSGDGMTVDIDKNTVPSDLTDSFTPIKIKNSGGEIIATGTVDNKNKQIYTT 300
Db 274 VTFEVDKVKSGDGMTVNIIDKNTVPSDLTDSFAIPKIKNSGGEIIATGTVDNTNKQIYTT 333

Qy 301 FTDVVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALSSVNKTIITVEYQRPENRT 360
Db 334 FTDVVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALSSVNKTIITVEYQRPENRT 393

Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDSTIIKVKYKVG 420
Db 394 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDSTIIKVKYKVG 453

Qy 421 DNQNLPSNRRIYDSEYEDVTNDYDIAQLGNNDVNIINFGNIDSPYIIKVISKYDPNKDDY 480
Db 454 DNQNLPSNRRIYDSEYEDVTNDYDIAQLGNNDVNIINFGNIDSPYIIKVISKYDPNKDDY 513

Qy 481 TTIQOVTMTQTTINEYGEFTASVDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 540
Db 514 TTIQOVTMTQTTINEYGEFTASVDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 573

Qy 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582
Db 574 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 615

RESULT 4
US-10-615-383-10
; Sequence 10, Application US/10615383
; Publication No. US20040038327A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS EPIDERMIDIS
; FILE REFERENCE: P06335US03/BAS
; CURRENT APPLICATION NUMBER: US/10/615,383
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-615-383-10

Query Match 93.3%; Score 2810.5; DB 12; Length 930;
Best Local Similarity 93.6%; Pred. No. 3e-164;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSINTDDNQ-IKKEETNNYDGIKRSERDTESTTNVDENEATFLOK 60

Db 73 SSNEEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 131

Qy 61 TPQDNTHLTETEEVKESSESSVSSSSIDTAQOPSHHTINREESVQTSNVDSDSHVDFPANS 120
Db 132 TPQDNTHLTETEEVKESSESSVSSSSMDTAQOPSHHTINSEASIQTSNDEENSRVDFPANS 191

Qy 121 KIKESNTESGKEENTIEQPNKVKEDSTTSOPSGYTNIDDKISNODELNLPINEYENKAR 180
Db 192 KIKESNTESGKEENTIEQPNKVRDSTTSOPSSYKNIDDKISNODELNLPINEYENKVR 251

Qy 181 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHLLIKVTDQSIETGYDDSEGVIIKAHDAENLIYD 240
Db 252 PLSTTSAQPSIKRVTVNQLAAEQSGSNVNHLLIKVTDQSIETGYDDSGI IKAHDAENLIYD 311

Qy 241 VTFEVDKVKSGDGMTVDIDKNTVPSDLTDSFTPIKIKNSGGEIIATGTVDNKNKQIYTT 300
Db 312 VTFEVDKVKSGDGMTVNIIDKNTVPSDLTDSFAIPKIKNSGGEIIATGTVDNTNKQIYTT 371

Qy 301 FTDVVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALSSVNKTIITVEYQRPENRT 360
Db 372 FTDVVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALSSVNKTIITVEYQRPENRT 431

Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDSTIIKVKYKVG 420
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSIIDSTIIKVKYKVG 491

Qy 421 DNQNLPSNRRIYDSEYEDVTNDYDIAQLGNNDVNIINFGNIDSPYIIKVISKYDPNKDDY 480
Db 492 DNQNLPSNRRIYDSEYEDVTNDYDIAQLGNNDVNIINFGNIDSPYIIKVISKYDPNKDDY 551

Qy 481 TTIQOVTMTQTTINEYGEFTASVDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 540
Db 552 TTIQOVTMTQTTINEYGEFTASVDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVDK 611

Qy 541 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 582
Db 612 DGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFDG 653

RESULT 5
US-10-689-082-10
; Sequence 10, Application US/10689082
; Publication No. US20040142348A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS EPIDERMIDIS
; FILE REFERENCE: P06335US04/BAS
; CURRENT APPLICATION NUMBER: US/10/689,082
; PRIOR FILING DATE: 2003-10-21
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-689-082-10

Query Match 93.3%; Score 2810.5; DB 16; Length 930;
Best Local Similarity 93.6%; Pred. No. 3e-164;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVNNQSINTDDNQ-IKKEETNNYDGIKRSERDTESTTNVDENEATFLOK 60

Db 73 SSNEEKNDVNNQSINTDDNQ-IKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 131

Qy 61 TPQDNTHLTETEEVKESSESSVSSSSIDTAQOPSHHTINREESVQTSNVDSDSHVDFPANS 120

Db 132 TPQDNTQLKEVVKEPSSVSSNSMDTAQPSHTTINSEASIQTSNEENSRVDFPANS 191
QY 121 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 180
Db 192 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKVR 251
QY 181 PLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 240
Db 252 PLSTTSAQPSKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSDGIIKAHDAENLIYD 311
QY 241 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFTIPKIKDNGSEIIATGTVDNKNQIITYT 300
Db 312 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFAIPKIKDNGSEIIATGTVDNKNQIITYT 371
QY 301 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 360
Db 372 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 431
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSIIDSTIIKVKYVG 420
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSIIDSTIIKVKYVG 491
QY 421 DNQNLPSNRIDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480
Db 492 DNQNLPSNRIDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 551
QY 481 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 540
Db 552 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 611
QY 541 DGIQNTNDNEKPLSNVLVTLYTPDGTSKSVRTDEGKYQFDG 582
Db 612 DGIQNTNDNEKPLSNVLVTLYTPDGTSKSVRTDEGKYQFDG 653

RESULT 6

US-10-690-184-10
; Sequence 10, Application US/10690184
; Publication No. US20040141997A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: METHODS FOR TREATING OR PREVENTING INFECTIONS FROM COAGULASE-
; FILE REFERENCE: P0633US05/BAS
; CURRENT APPLICATION NUMBER: US/10/690,184
; CURRENT FILING DATE: 2003-10-21
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-690-184-10

Query Match 93.3%; Score 2810.5; DB 16; Length 930;
Best Local Similarity 93.6%; Pred. No. 3e-164;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;
QY 1 SSDEEKNDVINNNQSINTDDNNQIIKKEETNNYDGIKESRDETRTESTTNVDENEATFLOK 60
Db 73 SSNEEKNDVINNNQSINTDDNNQIIKKEETNSDAIENRSKIDTQSTTNVDENEATFLOK 131
QY 61 TPQDNTQLKEEVKESSVSSNSIDTAQPSHTTINREESVQTSNDVDSHVSDFPANS 120
Db 132 TPQDNTQLKEEVKEPSSVSSNSMDTAQPSHTTINSEASIQTSNEENSRVDFPANS 191
QY 121 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 180

Db 192 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKVR 251
QY 181 PLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 240
Db 252 PLSTTSAQPSKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSDGIIKAHDAENLIYD 311
QY 241 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFTIPKIKDNGSEIIATGTVDNKNQIITYT 300
Db 312 VTFEVDKVKSGDVTVDIDKNTVPSDLTDSFAIPKIKDNGSEIIATGTVDNKNQIITYT 371
QY 301 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 360
Db 372 FTDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKLTALSSVNKTIITVEYQRPENRT 431
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSIIDSTIIKVKYVG 420
Db 432 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSIIDSTIIKVKYVG 491
QY 421 DNQNLPSNRIDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480
Db 492 DNQNLPSNRIDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 551
QY 481 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 540
Db 552 TTIOQTVMQTTINEYTGFEFTASVDNTIAFSTSSGQGGDLPEPKTYKIGDYVWEDVDK 611
QY 541 DGIQNTNDNEKPLSNVLVTLYTPDGTSKSVRTDEGKYQFDG 582
Db 612 DGIQNTNDNEKPLSNVLVTLYTPDGTSKSVRTDEGKYQFDG 653

RESULT 7

US-10-378-674-9
; Sequence 9, Application US/10378674
; Publication No. US2004006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-9

Query Match 88.6%; Score 2670; DB 15; Length 549;
Best Local Similarity 98.7%; Pred. No. 6.4e-156;
Matches 519; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDEEKNDVINNNQSINTDDNNQIIKKEETNNYDGIKESRDETRTESTTNVDENEATFLOK 60
Db 24 SSDEEKNDVINNNQSINTDDNNQIIKKEETNNNDGIEKSEDRTESTTNVDENEATFLOK 83
QY 61 TPQDNTQLKEEVKESSVSSNSIDTAQPSHTTINREESVQTSNDVDSHVSDFPANS 120
Db 84 SPQDNTLHTEEVKEPSSVSSNSIDTAQPSHTTINREESVQTSNDVDSHVSDFPANS 143
QY 121 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 180
Db 144 KIKESNTESKGEENTIEQPNKVKEDSTTSOPSGYTNIDEKISNODELLNLPINEYENKAR 203
QY 181 PLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 240
Db 204 PLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSITEGYDDSEGVIKAHDAENLIYD 263

Qy 241 VTPEVDDKVGSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTYDNKNKQIYTT 300
Db 264 VTPEVDDKVGSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTYDNKNKQIYTT 323
Qy 301 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPNNRT 360
Db 324 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPNNRT 383
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 420
Db 384 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 443
Qy 421 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480
Db 444 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 503
Qy 481 TTIIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 526
Db 504 TTIIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 549

RESULT 8
US-10-378-674-2
; Sequence 2, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-2

Query Match 83.3%; Score 2510.5; DB 15; Length 560;
Best Local Similarity 93.2%; Pred. No. 4.1e-146;
Matches 490; Conservative 14; Mismatches 21; Indels 1; Gaps 1;

Qy 1 SSDEKNDVINNQSINTDDNNQIIKKEETNNYDGIKRSSEDRTESTTNDENEATFLOK 60
Db 36 SSNEEKNDVINNQSINTDDNQ- IKKEETNSDAIENRSKDIQTOSTTNDENEATFLOK 94
Qy 61 TPQDNTHLTBEEVKSSSVSSSSIDTAQPSHTTTINREESVQTSNDVEDSHVDFANS 120
Db 95 TPQDNTQLKEEVKPESSVSSSSMDTAQPSHTTTINSEASIQTSNDNEENSRVDFANS 154
Qy 121 KIIESNTESKEENTIEQPNKVEDSTTSQPSGTYNTDEKISNODELLNLPINEYENKAR 180
Db 155 KIIESNTESKEENTIEQPNKVEDSTTSQPSGYNTDEKISNODELLNLPINEYENKVR 214
Qy 181 PLSTTSQAQPSKRVTVNQLAAEQSNVNHIIKVTQDSITTEGYDDSEGVKAHAENLIYD 240
Db 215 PLSTTSQAQPSKRVTVNQLAAEQSNVNHIIKVTQDSITTEGYDDSDGIIKAHAENLIYD 274
Qy 241 VTPEVDDKVGSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTYDNKNKQIYTT 300
Db 275 VTPEVDDKVGSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGEIIATGTYDNKNKQIYTT 334
Qy 301 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPNNRT 360
Db 335 FTDYVDKYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPNNRT 394
Qy 361 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 420
Db 395 ANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVG 454

Qy 421 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 480
Db 455 DNQNLPSNRIYDYSEYEDVTDNDYAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDY 514
Qy 481 TTIIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 526
Db 515 TTIIQQTVMQTINNEYTGERTASYDNTIAFSTSSGQGGDLPEK 560

RESULT 9
US-10-378-674-4
; Sequence 4, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-4

Query Match 55.5%; Score 1672; DB 15; Length 343;
Best Local Similarity 97.9%; Pred. No. 7.1e-95;
Matches 319; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 201 AEOGSNNHLIKVTDQSITEGYDDSEGVKAHAENLIYDVTFEVDDKVGSGDTMTVDID 260
Db 18 SEQGSNNHLIKVTDQSITEGYDDSGIILKAHAENLIYDVTFEVDDKVGSGDTMTVDID 77
Qy 261 KNTVPSDLTDSFTPIPKIKNSGEIIATGTYDNKNKQIYTTFTDYVDKYENIKAHKLKTSY 320
Db 78 KNTVPSDLTDSFALPKIKNSGEIIATGTYDNKNKQIYTTFTDYVDKYENIKAHKLKTSY 137
Qy 321 IDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPNNRTANLQSMFTNIDTKNHTVEQT 380
Db 138 IDKSKVPNNNTKLDVEYKLTALSSVKNKTIITVEYQRPNNRTANLQSMFTNIDTKNHTVEQT 197
Qy 381 IYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVDNQLPDSNRIYDYSEYEDV 440
Db 198 IYINPLRYSAKETNNVNSGNGDEGSTIIDSTIIKVKYKVDNQLPDSNRIYDYSEYEDV 257
Qy 441 TNDYDQAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIIQQTVMQTINNEYTGEF 500
Db 258 TNDYDQAQLGNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIIQQTVMQTINNEYTGEF 317
Qy 501 RTASYDNTIAFSTSSGQGGDLPEK 526
Db 318 RTASYDNTIAFSTSSGQGGDLPEK 343

RESULT 10
US-10-378-674-6
; Sequence 6, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/361,324
; PRIOR FILING DATE: 2002-03-05

```
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-6

Query Match      51.8%; Score 1562; DB 15; Length 316;
Best Local Similarity 97.7%; Pred. No. 3.7e-88;
Matches 298; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 201 AEOGSNNVHLIKVTQDSITEGVDSGVKKAHDAENLIYDVTFFVDDKVKSGDWTMTVDID 260
Db 12 SEQGSNNVHLIKVTQDSITEGVDSGVKKAHDAENLIYDVTFFVDDKVKSGDWTMTVDID 71
QY 261 KNTVPSDLTDSFTTPKIKDNGSGEIIATGTYDNKNKQIYTFYDVKYENIKAHKLKTSY 320
Db 72 KNTVPSDLTDSFAIPKIKDNGSGEIIATGTYDNKNKQIYTFYDVKYENIKAHKLKTSY 131
QY 321 IDKSKVPNNNTKLDVEYKTLSSVNTIIVYQRPENRNTANLQSMFTNIDTKNHTVEQT 380
Db 132 IDKSKVPNNNTKLDVEYKTLSSVNTIIVYQRPENRNTANLQSMFTNIDTKNHTVEQT 191
QY 391 IYINPLRYSAKETNNVNSGDEGSTIIDSTIIKVKYKVDNQNLPSNRIDYSEYEDV 440
Db 192 IYINPLRYSAKETNNVNSGDEGSTIIDSTIIKVKYKVDNQNLPSNRIDYSEYEDV 251
QY 441 TNDYDQAGNNNDVNIINFGNIDSPIIKVISKYDPNKDDYTTIOQVTMTTINEYTGFE 500
Db 252 TNDYDQAGNNNDVNIINFGNIDSPIIKVISKYDPNKDDYTTIOQVTMTTINEYTGFE 311
QY 501 RTASY 505
Db 312 RTASY 316

RESULT 11
US-10-282-122A-70251
; Sequence 70251, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUIRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70251
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-70251

Query Match      37.8%; Score 1137.5; DB 12; Length 1141;
Best Local Similarity 43.3%; Pred. No. 2.3e-61;
Matches 266; Conservative 97; Mismatches 193; Indels 59; Gaps 19;

QY 1 SSDEEKNDVINNQSDINTDDNQIIKKEETNNYDGIKESDRTSTNNVDENEATPLOK 60
Db 69 SDNKEVVSETEENSTTENDSTNPI--KKEINT--DSQPEAKEESTTSTQOQQNNVTATTE 125
QY 61 TPQDNTHLTFEEFVKES----SSVSSNSSIDTAQPSHTTINREESVQTSNVEDSHVSDF 117
Db 126 TKPQN--IEKENVKPSTKTATEDTSVILEEKAPNYT--NNDVTTKPS----- 170
QY 118 ANSKIKESNT---ESGKEENTIEQPNKVKEDSTTSOPSGVTNIDKIS--NODELLNLP-- 171
Db 171 -TSEIQTKPTTPOESTNIENSQPTPSKVD----NQVDTATNPKPEPVNVSKSELKNPEK 226
QY 172 -----INEYENKARPLST--TSAQF----STKRYTVNQALAAEOGSNNVHLIKVTQDSI 218
Db 227 LKELVRNDNNTDRSTKPVATAPTSPVAPKRLNACMRFAVAQPAAVASNNVNDLITVTKQTI 286
QY 219 TEGYDDSEGVKKAHDAENLIYDVTFFVDDKVKSGDWTMTVDIDKNTVPSDLTDSFTTPKIK 278
Db 287 KVG-DGKONVAAAHGDKGISEYDTEFTIDNKVKKGDTMTINYNKVNTPSOLTDKNDPIDIT 345
QY 279 DNSGEIIATGTYDNKNKQIYTFYDVKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYK 338
Db 346 DPSEVIAGTFDKATKQIYTFYDVKYEDIKARLTLYSIDKQAVP-NETSLMTFA 404
QY 339 TALSSVNTIIVYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVIS 398
Db 405 TAGKETSONVSDYQDPMVHGDSNIQSIFTKLDENKQITTEQIYVNPPLKKTATNTKVDIA 464
QY 399 GNG-----DEGSTIIDSTIIKVKYKVDNQNLPSNRIDYSEYEDVTDND-DYAO 448
Db 465 GSQVDDYGNIKLNGSTIIDQNTTEIKVKYKVPNQQLPQSNRIYDFSYQEDVTSQFQNKKS 524
QY 449 GNNNDVINFGNIDSPIIKVISKYDPNKDDYTTIOQVTMTTINEYTGFEFTASYDNT 508
Db 525 FSNVATLDFGDNISAYIIVKVSQYPTSDGELDIAQGTSMRTT-DKY-GYNYNAGYSNF 582
QY 509 IAPSTSSGGQGGDL-PPEKTYKIGDVYVWEDVDKQIGTQNTNDNEKPLSNVLTLTYPDGTS 567
Db 583 IVTSNDTGGGDTGTVKPEKLYKIGDVYVWEDVDKQVQGTDSKEKPMANVLTLTYPDGTT 642
QY 568 KSVRTDEDGKYOPDG 582
Db 643 KSVRTDANGHYEFGG 657

RESULT 12
US-10-282-122A-70444
; Sequence 70444, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 21.2079 Seconds
(without alignments)
2639.748 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SSDEKNDVINNNQSIINTDD.....PDGTSKSVRTDEBQKQFDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 T30214	fibrinogen-binding
2	1144.5	38.0	1166	2 T28680	fibrinogen-binding
3	1137.5	37.8	1141	2 E89824	hypothetical prote
4	576	19.1	1315	2 T28679	fibrinogen-binding
5	559	18.6	1385	2 D89824	hypothetical prote
6	489	16.2	953	2 C99824	hypothetical prote
7	452	15.0	933	2 S41539	fibrinogen-binding
8	452	15.0	989	2 D99852	fibrinogen-binding
9	397.5	13.2	877	2 P90070	Clumping factor B
10	376	12.5	940	2 S19702	fibronectin-bindin
11	344.5	11.4	961	2 G30053	hypothetical prote
12	326	10.8	1038	2 H90053	hypothetical prote
13	300.5	10.0	1018	2 A32192	fibronectin-bindin
14	239	7.9	1039	2 T30856	protein F2 - Strep
15	225	7.5	1463	2 T30280	AAS surface protei
16	216.5	7.2	1959	2 A31085	hypothetical prote
17	216.5	7.2	3394	2 T18501	hypothetical prote
18	212.5	7.1	1125	2 E90598	membrane nucleas
19	211.5	7.0	2206	2 G71611	hypothetical prote
20	208.5	6.9	1072	2 A86827	hypothetical prote
21	208	6.9	1127	2 T28317	ORF MSV156 hypothe
22	208	6.9	3724	2 T18427	hypothetical prote
23	207.5	6.9	1711	2 T18429	hypothetical prote
24	207.5	6.9	2401	2 T28676	rhodopy protein -
25	207	6.9	769	2 F89870	serine proteinase
26	207	6.9	1939	2 T18372	repeat organellar
27	207	6.9	4550	2 T18440	hypothetical prote
28	206.5	6.9	821	2 S67087	hypothetical prote
29	206	6.8	1302	1 JC6009	surface-located me

30	205.5	6.8	1365	2 T30822	Imp1 protein - Myc
31	205.5	6.8	2510	2 T28160	hypothetical prote
32	205	6.8	2481	2 D90011	FntB protein [mpo
33	204	6.8	3844	2 T18402	asparagine-asparta
34	203	6.7	537	2 A23770	asparagine-rich pr
35	202.5	6.7	1997	2 F71607	DNA helicase II BR
36	202.5	6.7	2269	2 T28677	rhodopy protein -
37	202	6.7	1093	2 T18275	1-phosphatidylinos
38	201.5	6.7	4688	2 F82885	hypothetical prote
39	199	6.6	665	2 B71609	hypothetical prote
40	198.5	6.6	807	2 B71605	hypothetical prote
41	198.5	6.6	1805	2 T02712	similar to late em
42	198.5	6.6	3147	2 T18674	hypothetical prote
43	198	6.6	1308	2 B71622	probable membrane
44	197	6.5	940	2 AD1374	internalin protein
45	197	6.5	2523	2 T18477	hypothetical prote

ALIGNMENTS

RESULT 1

T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:98261511; PMID:9596732
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NIL>
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 3013; DB 2; Length 1092;
Best Local Similarity. 100.0%; Pred. No. 7.2e-123;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SSDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSERDRTSTTNVDENATFLQK	60
Db	75	SSDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSERDRTSTTNVDENATFLQK	134
Qy	61	TPQDNTHLTBEVKESSESSNSSIDTAQPSHTTINREESVQTSNVEDSHVSDFANS	120
Db	135	TPQDNTHLTBEVKESSESSNSSIDTAQPSHTTINREESVQTSNVEDSHVSDFANS	194
Qy	121	KIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELINLPINYEYKAR	180
Db	195	KIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKISNQDELINLPINYEYKAR	254
Qy	181	PLSTSAQPSIKRVTVNQLAAEGSNVNHILIKVTDQSITEGYDDSEGVKAHDAENLIYD	240
Db	255	PLSTSAQPSIKRVTVNQLAAEGSNVNHILIKVTDQSITEGYDDSEGVKAHDAENLIYD	314
Qy	241	VTFEVDKVKSGDVTMTVDIDKNTVPDLTDSFTIPKIKDMSGELIATGTDNKNKQIYTT	300
Db	315	VTFEVDKVKSGDVTMTVDIDKNTVPDLTDSFTIPKIKDMSGELIATGTDNKNKQIYTT	374
Qy	301	FTDYVDKYENIKAHKLKTSYIDKSQVNNNTKLDVEYKLTALSSVNKTIITVEYQPNENRT	360
Db	375	FTDYVDKYENIKAHKLKTSYIDKSQVNNNTKLDVEYKLTALSSVNKTIITVEYQPNENRT	434
Qy	361	ANLQSMFTNIDTKNHTVEQTIYINPLYSAKETNNVNSGNGDEGSTIIDDSITIKVKYVG	420
Db	435	ANLQSMFTNIDTKNHTVEQTIYINPLYSAKETNNVNSGNGDEGSTIIDDSITIKVKYVG	494
Qy	421	DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNFGNIDSPYIIKVISKYPNKKDDY	480
Db	495	DNQNLPSNRIYDYSEYEDVTNDDYAQLGNNDNVNFGNIDSPYIIKVISKYPNKKDDY	554

RESULT 4

T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C/Species: Staphylococcus aureus
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C/Accession: T28679
R/Josefsson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.; Microbiology 144, 3387-3395, 1998
A/Title: Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus
A/Reference number: Z20510; MUID:99098700; PMID:9884231
A/Accession: T28679
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1315 <JOS>
A/Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA06651.1
C/Genetics:
A/Gene: sdrD

```
Query Match      19.1%; Score 576; DB 2; Length 1315;
Best Local Similarity 30.1%; Pred. No. 1.5e-17;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

Qy      1 SSDEEKNDV---INNQSINTDDNNQIIKKEETNNYDGIKRSSEDRTESTNVDS-NEAT 56
Db      1 SSDEEKNDV---INNQSINTDDNNQIIKKEETNNYDGIKRSSEDRTESTNVDS-NEAT 56
Qy      55 STNKELNEATTSASDNQSDKVDQMQLNQEDNTKNDQKEMVSSQGNHTTSNGNKLSIEKE 114
Db      55 STNKELNEATTSASDNQSDKVDQMQLNQEDNTKNDQKEMVSSQGNHTTSNGNKLSIEKE 114
Qy      57 FLOKTPQDNTHLTBEVEKSESSVSSSSIDTAQPSHTTINREESVQTSQDNVEDSHVSD 116
Db      57 FLOKTPQDNTHLTBEVEKSESSVSSSSIDTAQPSHTTINREESVQTSQDNVEDSHVSD 116
Qy      115 SVQSTTGKNEVSTAKSDEQSPKSTNEDLNTQ-----TISNQELQ-PDLQENKSVVN 168
Db      115 SVQSTTGKNEVSTAKSDEQSPKSTNEDLNTQ-----TISNQELQ-PDLQENKSVVN 168
Qy      117 FANSKIKESNTSGKEENTIEQPN-KVKEDSTTSQPSGYTNI-DEKISNQDELL--NLPIN 173
Db      117 FANSKIKESNTSGKEENTIEQPN-KVKEDSTTSQPSGYTNI-DEKISNQDELL--NLPIN 173
Qy      169 -----VQPTNEENKVD-----AKTESTT-----LNKSDAIKSENDETLVDNNSNS 209
Db      169 -----VQPTNEENKVD-----AKTESTT-----LNKSDAIKSENDETLVDNNSNS 209
Qy      174 EYENKARPLSTTSQAP-----SIKRVTVNQLAAEQSNVHLIKVTDQSIETEGYDDSEGI 229
Db      174 EYENKARPLSTTSQAP-----SIKRVTVNQLAAEQSNVHLIKVTDQSIETEGYDDSEGI 229
Qy      210 NNENNADIIIPKSTAPKELNTRMRAIAVQPSSTEAQVNDLITNTTLTVVDADKNKIV 269
Db      210 NNENNADIIIPKSTAPKELNTRMRAIAVQPSSTEAQVNDLITNTTLTVVDADKNKIV 269
Qy      230 KAHAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTPIKID-NSG 282
Db      230 KAHAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTPIKID-NSG 282
Qy      270 PAQDYLKLSQIT--VDDKVKSGDYFTIKY--SDTVQVYGLNPEDIKN---IGDIKDPNNG 323
Db      270 PAQDYLKLSQIT--VDDKVKSGDYFTIKY--SDTVQVYGLNPEDIKN---IGDIKDPNNG 323
Qy      283 ELIATGTVDNKNQIYTFDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKALS 342
Db      283 ELIATGTVDNKNQIYTFDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKALS 342
Qy      324 ETIATAKHDTANLLITYTFDYVDFRNSVGMGINSYIYMDADTIP--VSKNDVEFNVTIG 381
Db      324 ETIATAKHDTANLLITYTFDYVDFRNSVGMGINSYIYMDADTIP--VSKNDVEFNVTIG 381
Qy      343 SWNKITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEOTIYINPLRYSAKE 392
Db      343 SWNKITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEOTIYINPLRYSAKE 392
Qy      382 NITTKYANIQYPDVYVNEKNSIGSAFT--ETVSHVGNKENPGYKQTIYVNPSENLTN 439
Db      382 NITTKYANIQYPDVYVNEKNSIGSAFT--ETVSHVGNKENPGYKQTIYVNPSENLTN 439
Qy      393 TNVNI-----SGNGDEGSTIIDSTIIKVKYVGDNQNLPSNRIDYS--EYEDVTNDIYA 446
Db      393 TNVNI-----SGNGDEGSTIIDSTIIKVKYVGDNQNLPSNRIDYS--EYEDVTNDIYA 446
Qy      440 AKLVQVAVHSSVPNNIGQINKDVTDIKIYQVPKGYTL--NKGYDVNTKELTDTVTN-QYL 495
Db      440 AKLVQVAVHSSVPNNIGQINKDVTDIKIYQVPKGYTL--NKGYDVNTKELTDTVTN-QYL 495
Qy      447 Q---LGNNDVNIINFGNIDSPYIIKVISKYPDNKDDYTTIIQQTVMQTINNEYTGEFRTA 503
Db      447 Q---LGNNDVNIINFGNIDSPYIIKVISKYPDNKDDYTTIIQQTVMQTINNEYTGEFRTA 503
Qy      496 QKITYGDNNSAVIDFGNADSAVYVNVNTKFOYTNSSEPTLVQMATLSSTGN-----KSV 549
Db      496 QKITYGDNNSAVIDFGNADSAVYVNVNTKFOYTNSSEPTLVQMATLSSTGN-----KSV 549
Qy      504 SYDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVKDQIGNTNDNEKPLSNVLTLTYTP 563
Db      504 SYDNTIAPSTSSGQGGDLPEKTYKIGDYVWEDVKDQIGNTNDNEKPLSNVLTLTYTP 563
Qy      550 STGNALGFTNQSGAG-----QEVYKIGNYVWEDTNKNGVQEL--GEKGVGNVTVTV-FD 602
Db      550 STGNALGFTNQSGAG-----QEVYKIGNYVWEDTNKNGVQEL--GEKGVGNVTVTV-FD 602
Qy      564 DGTSKSVR---TDDEGKY 578
Db      564 DGTSKSVR---TDDEGKY 578
Qy      603 NNTNTKVGAEVTKEDGSY 620
Db      603 NNTNTKVGAEVTKEDGSY 620
```

RESULT 5

D89824
hypothetical protein sdrD [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: D89824

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc-
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: D89824
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1385 <KUR>
A/Cross-references: GB:BA000018; PID:g13700454; PIDN:BA041751.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: sdrD

```
Query Match      18.6%; Score 559; DB 2; Length 1385;
Best Local Similarity 28.8%; Pred. No. 8.5e-17;
Matches 177; Conservative 117; Mismatches 235; Indels 86; Gaps 27;

Qy      1 SSDEEKNDV---INNQSINTDDNNQIIKKEETNNYDGIKRSSEDRTESTNVDS-EAT 56
Db      1 SSDEEKNDV---INNQSINTDDNNQIIKKEETNNYDGIKRSSEDRTESTNVDS-EAT 56
Qy      55 STNKELNEATTSASDNQSDKVDQMQLNQEDNTKNDQKEMVSSQGNHTTSNGNKLSIEKE 114
Db      55 STNKELNEATTSASDNQSDKVDQMQLNQEDNTKNDQKEMVSSQGNHTTSNGNKLSIEKE 114
Qy      57 FLOKTPQDNTHLTBEVEKSESSVSSSSIDTAQPSHTTINREESVQTSQDNVEDSHVSD 116
Db      57 FLOKTPQDNTHLTBEVEKSESSVSSSSIDTAQPSHTTINREESVQTSQDNVEDSHVSD 116
Qy      115 SVQSTTGKNEVSTAKSDEQSPKSTNEDLNTQ-----TISNQEGIQ-PDLLENKSVVN 168
Db      115 SVQSTTGKNEVSTAKSDEQSPKSTNEDLNTQ-----TISNQEGIQ-PDLLENKSVVN 168
Qy      117 FANSKIKESNTSGKEENTIEQPN-KVKEDSTTSQPSGYTNI-DEKISNQDELLNLPINEY 175
Db      117 FANSKIKESNTSGKEENTIEQPN-KVKEDSTTSQPSGYTNI-DEKISNQDELLNLPINEY 175
Qy      169 -----VQPTNEENKVDAKTESTTTLNVKSDAIKS-----NAETLVNNSNSNN----- 211
Db      169 -----VQPTNEENKVDAKTESTTTLNVKSDAIKS-----NAETLVNNSNSNN----- 211
Qy      176 ENKARPLSTTSQAPSIKRVTVNQLAAEQS-----NVHLIKVTDQSIETEGYDDSEGIKA 231
Db      176 ENKARPLSTTSQAPSIKRVTVNQLAAEQS-----NVHLIKVTDQSIETEGYDDSEGIKA 231
Qy      212 ENNADIIIPKSTAPKSLNTRMRAAIQPNSTDSKNVNDLITNTTLTVVDADNSKTIVPA 271
Db      212 ENNADIIIPKSTAPKSLNTRMRAAIQPNSTDSKNVNDLITNTTLTVVDADNSKTIVPA 271
Qy      232 HAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTPIKID-NSGEI 284
Db      232 HAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTPIKID-NSGEI 284
Qy      272 QDYLKLSQIT--VDDKVKSGDYFTIKY--SDTVQVYGLNPEDIKN---IGDIKDPNNGET 325
Db      272 QDYLKLSQIT--VDDKVKSGDYFTIKY--SDTVQVYGLNPEDIKN---IGDIKDPNNGET 325
Qy      285 IATGTVDNKNQIYTFDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKALS 344
Db      285 IATGTVDNKNQIYTFDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYKALS 344
Qy      326 IATAKHDTANLLITYTFDYVDFRNSVGMGINSYIYMDADTIPVD--KKDVPFVSFTIGNQ 383
Db      326 IATAKHDTANLLITYTFDYVDFRNSVGMGINSYIYMDADTIPVD--KKDVPFVSFTIGNQ 383
Qy      345 NKITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEOTIYINPLRYSAKETN 394
Db      345 NKITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEOTIYINPLRYSAKETN 394
Qy      384 ITTTADITTPAYKADNNSIGSAFT--ETVSHVGNVEDFGYVNVVYVNPMDKLGAK 441
Db      384 ITTTADITTPAYKADNNSIGSAFT--ETVSHVGNVEDFGYVNVVYVNPMDKLGAK 441
Qy      395 VNISNGDEGST---IIDSTIIKVKYVGDNQNLPSNRIDY--SEYEDVTND--DYA 446
Db      395 VNISNGDEGST---IIDSTIIKVKYVGDNQNLPSNRIDY--SEYEDVTND--DYA 446
Qy      442 LKVEAYHPKPTNIGQINQVNTNIIKIRVPEGYTL---NKGYDVNTDLVDVDFEKNKM 498
Db      442 LKVEAYHPKPTNIGQINQVNTNIIKIRVPEGYTL---NKGYDVNTDLVDVDFEKNKM 498
Qy      447 QLGNNDVNIINFGNIDSPYIIKVISKYPDNKDDYTTIIQQTVMQTINNEYTGEFRTASYD 506
Db      447 QLGNNDVNIINFGNIDSPYIIKVISKYPDNKDDYTTIIQQTVMQTINNEYTGEFRTASYD 506
Qy      499 TYGSNQSVNLDFDGTISAYVNVNTKFOYTNSSEPTLVQMATLSSTGN-----KSVSTG 552
Db      499 TYGSNQSVNLDFDGTISAYVNVNTKFOYTNSSEPTLVQMATLSSTGN-----KSVSTG 552
Qy      507 NTAFTSSGQGGDLPEKTYKIGDYVWEDVKDQIGNTNDNEKPLSNVLTLTYPDGT 566
Db      507 NTAFTSSGQGGDLPEKTYKIGDYVWEDVKDQIGNTNDNEKPLSNVLTLTYPDGT 566
Qy      553 NALGFTNQSGAG-----QEVYKIGNYVWEDTNKNGVQEL--GEKGVGNVTVTV-FDNNT 605
Db      553 NALGFTNQSGAG-----QEVYKIGNYVWEDTNKNGVQEL--GEKGVGNVTVTV-FDNNT 605
Qy      567 SKSVR---TDDEGKY 578
Db      567 SKSVR---TDDEGKY 578
Qy      606 NTRKVGAEVTKEDGSY 620
Db      606 NTRKVGAEVTKEDGSY 620
```

RESULT 6

C89824
hypothetical protein sdrC [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: C89824
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc-
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: C89824
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-953 <RUR>
A/Cross-references: GB:BA000018; PID:G13700453; PIDN:BA841750.1; GSPDB:GN00149
A/Experimental source: strain N315
A/Genetics: C/Genetics:
A/Gene: sdrC

Query Match	16.2%	Score 489;	DB 2;	Length 953;
Best Local Similarity	28.2%	Pred. No. 5.5e-14;		
Matches 162;	Conservative	82;	Mismatches 236;	Indels 94;
				Gaps 22;

35	QY	35	GLEKSGEDTEGTTNVDBNEATFLQKTPQDNHILTEBEVKSESSVSSNSIDTAQPSSH	94
45	Db	45	GHEAKAAEHTNGELNQSNETT---APSENK--TTEKV--DSROLKDNQTATAQPKV	96
95	QY	95	TTINREESVQTSNDVSDSHVDFPANSKIKESNTESGKEENTIEOPNKKVEDSTISPSGY	154
97	Db	97	T-----MDSATVKETSSNMOS-----PONATASQSTTQTTSTNV	129
155	QY	155	TNIDEX---ISNQDELLNLPINEVEN--KARPLSTTSAQPSIKRVTVNOLAA--EOGSNVNH	209
130	Db	130	TTNDKSSSTYSNETDKSNL--TOAKNVSTTPTKTTIIKORALNRWNVTVAAPQOGTVND	187
210	QY	210	LKVTQDSI-----TEGYDDEGVIKAHDAENLIYDVFEDVDKVKSGDTMT	256
188	Db	188	KVHFTNIDIAIDKGHVNKTGTGTEFWATSDVLK-----LKANYTIDDSVKEGUTFT	239
257	QY	257	VDIDKNTVPSDLTDSFTPIKIDNSGETIATGTVDNKNKQITVTFDQYDKYENIKAHUK	316
240	Db	240	FKYQVFRPGSVRLPSQFONLYNAQGNIIAKGIYDSKNTTNTFTNYDQYTVNVSSE	299
317	QY	317	LASYIDKSKVPNNNTKLDVBYKTALSSVNKTIITVEYQRPNEENTANLOSMFTNIDFKNHT	376
300	Db	300	QVAFAKRENATDKTAYEMEVTLGNDYYSKDIVDY---GNQKGQQLIISTNYINNEDS	356
377	QY	377	VEQTYIYN--PURYSAKETNV--NISGNGDEGSTIIDDSIIKVKYKVGDNQNLNPDNSRIYD	434
357	Db	357	RNMTVYVNPQPKTYTKETFTVNLN-----GYKENPDAKNFKIYEVTDOQFVDS--FTPDT	410
435	QY	435	SEVEDVTND-DYAQLGNNDVYNIFGN-----IDSPYIIKVIKSYDKNKDDYTIIOQTVM	489
411	Db	411	SKGLDVTQGFVIVSYNDMKNKATVDLLNGQSSDKQYIIQQVAYPDSSTYDNGIKDYTLET	470
490	QY	490	QTTINEXTGEFRTASYNDTIAFSTSSGOGGDLPPKTYKIGYVWVEDVDKGIQNTNDN	549
471	Db	471	QNGKSSWEN-----SYSNVNGSSSTANGD-----QKYNLGDYVWEDTNKDGKQDA--N	516
550	QY	550	EKPLSNVLVTLYTPDGTG--KSVRTDEBGKYQFDG	582
517	Db	517	EKIGKGVTVILKDNKGKELDRTTTTDENGKQFTG	550

RESULT 7
S41539
fibrinogen-binding protein - *Staphylococcus aureus*
N:Alternate names: clumping factor
C:Species: *Staphylococcus aureus*
C:Date: 13-Jan-1995 #sequence_revision
C:Accession: S41539; S36630
R:McDevitt, D.; Francois, P.; Vaudeaux, P.; Foster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*
A:Reference number: S41539; PMID:94224142; PMID:8170386
A:Text: 13-Jan-1995 #text_change 15-Oct-1999

A/Cross-references: EMBL:Z18852; NID:g397525; PID:CAA79304.1; PID:g397526
 A/Rebukes: 1-933 <MCD>

Query Match	15.0%;	Score 452;	DB 2;	Length 933;
Best Local Similarity	26.6%;	Pred. No. 2.1e-12;		
Matches 154;	Conservative 105;	Mismatch 234;	Indels 86;	Gaps 26;
Qy	1	SSDEKNDVNNQGSINTDNNQIIKKEETNNYDGIETKRSDETRSTNTVDNEATPQK 60		
Db	51	SNESKNSDSSVSAAPKTDITNV-----SDTTSSTNN-----NGETSAVQ 91		
Qy	61	TPQDNTHLTETEEVKSESSVSSNSIDTAQOPSHHTINREESVQTSNDVSDHVSDFPANS 120		
Db	92	NP-----AQOETTCSSSTWATTEETPVTCGEATTTTINQANTPATQCS-SNTWAEELVN- 143		
Qy	121	KIKESNTESKEENTIEOPKNVKEDSTTSOPSGYTNIDEXIKNODELLN-LPINEYENKA 179		
Db	144	--QTSNETTTNDTNTV-----SSVNSPQNSTNAENVSTTQDTSPEATPSN---NES 189		
Qy	180	RPLSTTSAQPSIKRVTNV-----QLAAEQGSNNHLIKVTDQ--SITEGVDDSEGV 228		
Db	190	APQSTDASNDVVNQAVNTSAPRMRAFSLAAVAADAPAGTDITNQTUNTVGL-DSGTT 248		
Qy	229	IKAHDAENLIYDVTFEVDKVKSGDTMTVIDKNTVPSDLTDSFTPIKIDNSGEIATG 288		
Db	249	VYFHOAGYVKLNTGFSVPNSAVKGDFTKIVPKELNLNGVTSTAKVPPIMAGD-QVLANG 307		
Qy	289	TYDNKNQIITYTDDYVDYKVENIKAHKLTSYTDKSKVPNNNTKL-DVEYKTALSS--VN 345		
Db	308	VIDS-DGNVITYTDDYVNTKDDKATLTMPAYID----PENVKKTGNVTLATIGISTTAN 362		
Qy	346	KTITVEQRENERTANLOSMTFNIDTKNHTVEQTIYINPL--RYSAKETNNVNISNGDE 403		
Db	363	KTVLVDYKYGKGFVNLISIKGTIDQIDKTNNTYHQTIVVNFSGDNLVAPVLTCNLKPNTDS 422		
Qy	404	GSTIIDSTYIIVKYVGDGNQNLPSNRRIYDYSEYEDVTDNDDDYAQLGNNDVNINFG---- 459		
Db	423	NALIDQOOTSIKYKVDNAADLSESYFV-NPENFEDVTNSVNITTFPNPQYKVEFTNPDD 481		
Qy	460	NIDSPYIIVKISKYDPN-KDDYTTIQOVTVMOTTINEYTG--FRTASVDNTIAFTSSG 516		
Db	482	QITTPYIVVNGHIDFNSKGD-----LALASTLYGYSNIIWRSMNSDNEVAFNNGSG 534		
Qy	517	QGGG-DLP--PEKTYKIGY--VWEDVDKD-GIQNTNDN 549		
Db	535	SGGDIDKPVVPEOPDPFGIEPIPEDSDSDPGSDSGSDS 573		

RESULT 8
D89852
fibrinogen-binding protein A, clumping factor [imported] - *Staphylococcus aureus* (strain
C/Species: *Staphylococcus aureus*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: D89852
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc-
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Ma-
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Ritle: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A/Reference number: A89758, PMID:21311952; PMID:11418146
A/Accession: D89852
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-989 <KUR>
A/Cross-references: GB:BA000018; PID:g13700678; PIDN:BA041975.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: clfa

```

Query Match      15.0%; Score 452; DB 2; Length 989;
Best Local Similarity 25.9%; Pred. No. 2.3e-12;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

QY      1  SDDEEKNNVNNQNSINTDDNNOI KKEETNNYDGIKESDTESTTNNVDNEATFLQK 60
      |: : : : : |: : |: : |: : |: : |: : |: : |: : |: : |: :

```

```
Db 51 SNESKNDSSVSRAPKTDDTNV-----SDTKSSNTN---NGETSVAQ 91
Qy 61 TPQDNTHLTBEVKESSVESNSSIDTAQPSHTTINREESVQTSNDVEDSHVDFPANS 120
Db 92 NP-----AQETTSSTNATETPTVTEATTTTNTQANTPATQS-SNTNAEELVN- 143
Qy 121 KIKESNTESGKEENTIEOPKNVKEDSTTSQPSGVTNIDEKISNODELLN-LPINEYENKA 179
Db 144 ---QTSNETTNDTNTV-----SSVNSPONSNAENVSTTQDTSTEATPSN---NES 189
Qy 180 RPLSTISA-----QPSIKRVTVNQLAA-----EAGSNVNHLIKVTDOISITEGVDD 224
Db 190 APQNTDASKNVDVQAVNPSTPRRAFSLAADAAPAGTDITN--QLTDVKVT---ID 244
Qy 225 SEGVIKAHAENLIYDVTFFVDDKVKGDVTMTVDIDKNTVPSDLTDSFTPIKIKDNSEGI 284
Db 245 SGTTVYHOAGVYKLVNGFSPNSAVKGDTEKFTVPKELNLNGVSTAKVPPIMAGD-QV 303
Qy 285 IATCTYNNKQIYTFDVTVDYVKYENIKAHKLTSYDKSKVPNNKVL-DVEYKTL-- 341
Db 304 LANGVIDS-DGNVITYTFDYVDNKENVTANITMPAYID----PENVTKTGNVTLTTGIGT 358
Qy 342 SSVAKTITVEYQRPNEKNTALQSMFTNIDTKNHTVEQTIYINLPRYSAKETNV----- 395
Db 359 NTASKTVLIDYEKGQFNLISIKGTIDQIDKNTNYRQTIYVNP-----SGDNVVLPAIT 413
Qy 396 -NISGNDGEGSTIIDSTIIKVKYKVDNQNLPSDNRIYDYSEYEDVTDNDYVAQLGNNDV 454
Db 414 GNLPNTKSNALIDAKNTDIIKVRV-DWANDLSSYYVNPSPDFEDVTNQVRISPPNAYQ 472
Qy 455 NINF-----GNTDSPYIIKISKYDPNKDDYTIQOTVTMTQTIINEYGEF--RTASVDNT 508
Db 473 KVEFPDQDDQITPYIVVNGHIDPASTG-----DLALRSTFYGDSNFWIRSMWDNE 526
Qy 509 IAFSTSSQGGQ-DLP--PEKTYKIGDY--VVEDVDKD-GTQNTNDN 549
Db 527 VAFNNGSGGQIDKPVVPEQDFPEGEIEPIPEDSDPDSGSDS 573

RESULT 9
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUP>
A:Cross-references: GB:BA000018; PID:g13702588; PIDN:BA843728.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 13.2%; Score 397.5; DB 2; Length 877;
Best Local Similarity 24.5%; Pred. No. 4.4e-10;
Matches 143; Conservative 81; Mismatches 207; Indels 153; Gaps 23;

Qy 2 SDEKNDVI---NNQSTINDNNQI IKKETNNYDGIKESEDRTESTTNNVDNEATFL 58
Db 44 ASEQSDNTQSSKNASADSEKNNMI-----ETPQLTNTANDTSDISANTNSANVDSTTKP 99
Qy 59 QKTPODNTHLTBEVKESSVESNSSIDTAQPSHTTINREESVQTSNDVEDSHVSDFA 118
Db 100 MSTQTSNTTTE-----PASTNETPQTAIKNQ---ATAAKQDQTVQEA 142
Qy 119 NSKIKESNTESGKEENTIEOPKNVKEDSTTSQPSGVTNIDEKISNODELLNLPINEYENK 178
```

```
Db 143 NSQVDNKT---NDANSIATNSLKNQST-----LDLP-----Q 173
Qy 179 ARPLSTTSAQ-----PSIKRVTVNQLAAEQ-----GSNVNHLIKVTDOISITEGVDD 225
Db 174 SSPQITISNAQGTSPKSPVRTRAVSLA VAEPPVNAADAKGTNVN-----DKVTASNFKLE 227
Qy 226 EGVIKAHAENLIYDVTFFVDDKVKSGDTMTV-----DID---KNTVPSDLTDS 271
Db 228 KTTTFDNPQSGNTFMAANFTVTDKVKSIGDYFAKLPSLDTGNGDVDVSNSTNP----- 281
Qy 272 FTIPKIKDNSEGIATCTYNNKQIYTFDVTVDYVKYENIKAHKLTSYDKSKVPNNNT 331
Db 282 --IADIKSTNGDVVAKATYDILTKTYTFVFDVNNKENINGQFSLPLFDRAKAPKSGT 339
Qy 332 KLDVEYKALTSSVKNKTIITVEYQRP-----NENRTANLQSMFTNIDTKN--HTVEQTIYINP 385
Db 340 -YDANINIADEMFPNNKITNYSPFIAGIDKPGANISSQIIGVDTASGQNTYKQTVFVNP 398
Qy 386 LRYSAKETNVNISGNDGEGSTIIDDS-----TIKVKYKVDNQNLPS----- 428
Db 399 KORVLGNTWYIIRGYQDK-----IESSGKVSATDTKLRIFEVNDTSKLSDSYADPNDSN 454
Qy 429 -----NRIYDYSYEDVTNDYVAQLGNNDVNINFGNIDSPYIIKISKYDPNKDD 479
Db 455 LKEVTDQFKRIY-----YE-----HPNVASIKFGDITKYTVVLVEGHYDNTGKN 499
Qy 480 YTT--IQOTVTMTQTIINEYGEFRTASYDNTIAFSTSSQGGQGD 521
Db 500 LKQTVIQENVDPVT-----NRDYSIFGNNNENVVRYGGGSADGD 538

RESULT 10
S19702
fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S19702
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus
A:Reference number: S19702; MUID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <JOE>
A:Cross-references: EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g581562
C:Keywords: fibronectin binding

Query Match 12.5%; Score 376; DB 2; Length 940;
Best Local Similarity 23.4%; Pred. No. 4.1e-09;
Matches 138; Conservative 110; Mismatches 217; Indels 124; Gaps 22;

Qy 35 GIEKRSSEDRTESTNVNDENATFLQKTPQDNTHLTBEVKESSVESNSSIDTAQ--- 90
Db 30 GQEKAAASEQNNTTVEES-----GSSATESKASSETQTNTNNVTIDTQSYSA 78
Qy 91 ----QPSHTT-INREESVQT--SDNVEDSHVDFPANSKIKESNTESGKEENTIEOPKNVK 143
Db 79 TSTEQPSQSTQVTTTEAPKTVQAPKVSrv-DLPSEKVDKETTGTQVD--IAQPSNV 135
Qy 144 EDSTTSQPSGVTNIDEKISNODELLNLPINEYENKAPLSTTSQAQPSIKRVTVNQLAAEQ 203
Db 136 E-----IKPRMKGSTVDVTAEEK 153
Qy 204 -----GSNVNHLIKVTDOISITEGVDDSEGVIKAHDAENLIYDVTFFVDDKVKSGDT 254
Db 154 EVVEETKATGTDVTKNKEVEGSEIGVGHKQDNTNVNPNHNAERVTLKVKFKPGEIGKAGDY 213
Qy 255 MTVDIDKNTVPSDLTDSFTIPKIKDNSEGIATCTYNNKQIYTFDVTVDYVKYENIKAH 314
Db 214 FDFLSDNVETHGISTLRKVPKISTDQVWATGEIIGERK-VRYTFKEYVQEKDLTAE 272
```



```
QY 315 LKLTSDVSKVPNNNTKLDVEYKTSALSSVNTKIIVFY---QRPNNRTANLQSMFTNID 371
DB 273 LSLNLFIDFTTQKNGO--NVEYKLGCTTFSKIFNIQLGGVRDNGWGTAN-----GRID 326
QY 372 TKNHT---VEQITLYINPLRYSAKETNVAISNGDSEGTIIDSTIIKVKYKCDNQLPDS 428
DB 327 TLNKVDGKSFHAYMKPNQSL--SSVTVTGVTGKNGKPGVNNPTVKYKHIGSDDLAEB 384
QY 429 --NRIYDSEYEDVTDNDYDQAQGNNDNVNINFGNIDSP--YIIKVISKYDPNKKDYTTIQ 484
DB 385 VYAKLDDVSKFEDVTDNNSLDFDTGGYSLNENNLQSKNYVIKGYDYSNASN-----439
QY 485 QVTVMTQTTINETGFERIA--SYDNTIAPSTSGGQGG-----DLPPEKTKIG 531
DB 440 --LEFQTHLFGYNYIYTTSLNLTWKGVAFYNSNAQDGKDKLKEPIIEHSPIELEPKSE 497
QY 532 DVVWEDVDKDGIONTNDNEKPLSNVLTLT-----YPDGTSKSVPTDSD 575
DB 498 PPVEKHELTGTLIESNDS--KPIDFHYHTAVEGAEGHAGT---IETED 542

RESULT 11
G90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PIDN:BAB43593.1; GSPDB:GN00149
A:Accession: G90053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702452; PIDN:BAB43593.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 11.4%; Score 344.5; DB 2; Length 961;
Best Local Similarity 22.7%; Pred. No. 9.6e-08;
Matches 120; Conservative 125; Mismatches 203; Indels 81; Gaps 22;

QY 74 KESSVESNSISDPAQPSHTTINREESVOTSDNVEDSHVSDFANSKIKESNTESGKEE 133
DB 33 KEAASEQNNTVBSGS-SATESKASETQTTNNV-----NTIDETQSYSA--- 78
QY 134 NTIEQPNKVKEDSTTSQPSGY--TNIDEKISQDELNLNLPINYEYK-----ARPLSTT 185
DB 79 TSTEQPSKSTQVTTBEAPTVAQPKVETEMSKSQEDLPSEKVADEKTTGTQVDIAQPNVS 138
QY 186 SAQPSIKE-----VTVNQLAAE---QGSNVNHLIKVTDQISITEGYDDSEGVIAKHAEN 236
DB 139 EIKPMKKSADVTASEKEAAEKATGTDVTKNKEVTESSL--EGHNKDSNVNPNHNAQR 197
QY 237 LIYDVTFEVDKVKSGDWTMTVDIKNTVPDLTDSFTIPKIDNS--GEIATGTVDNKNKQIT 295
DB 198 VTLKMKFGEGIRAGDYFDFTLSDNVETHGISTLRKVPKESSTEDKMWANGQVINE-R 256
QY 236 QITVFTDYDYKYENIKAHKLKTSYIDKSQVNNNTKLDVEYKTSALSSVNTKIIVFYORP 355
DB 257 TIRYFTFDYINNKDGLTAEALNLFIDPTTQKGSQ--KVEVTLGQNKVSKFEIDIKYLDG 315
QY 356 NENR---TANLQSMFTNIDTKNH---TVEQITLYINPLRYSAKETNVAISNGDSEGTIID 409
DB 316 VKDRMGVTN-----GRIDTLNKEGKSFHAYMKPNQSL--TSVTVTGVTSGYKQSA 368
QY 410 DSTIIKVKYKCDNQLPDS--NRIYDSEYEDVTDNDYDQAQGNNDNVNINFGNIDS--PY 465
DB 369 NNPTVKYKHIGSDELAESVYAKLDDTSKPEDVTEKVNLSYTSNGGYTLNIGLDLNSKDY 428
```

```
QY 466 IIKVISKYDPNKKDYTTIQOTVTMTQTTINETGEBRTASY-----DNTIAPSTSS 515
DB 429 VIKYEGEDQNAKD-----LNFRTHLSGHVKYPPYPPYPPYVQLTWNNGVAFYSNN 481
QY 516 GQGGG-DLPPEKTKYKIDGYVWEDVDKD-----GIQNTNDNEKPL 553
DB 482 AKGDKPKNPDIIEKSEPIDLDIKSEPPVVEKHELTGTLIESNDS-KPI 529

RESULT 12
H90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PIDN:BAB43594.1; GSPDB:GN00149
A:Accession: H90053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702453; PIDN:BAB43594.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 10.8%; Score 326; DB 2; Length 1038;
Best Local Similarity 22.7%; Pred. No. 6.7e-07;
Matches 137; Conservative 109; Mismatches 235; Indels 122; Gaps 30;

QY 35 GIEKRSDETESTNVDENEATFLQKTPQDN----THLTFEEYKESSESVSSSSSIDTAQ 90
DB 30 GODKEAASQKTTVEEN-----GNSATDNKTSQTATTNVNHHIEETOSYNATV--TE 82
QY 91 QPSHTT-INREESVOTSDNVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVEDSTTS 149
DB 83 QPNAQTQVTEEPKAVQ-----APQTAQPAVETVKEE---EKP-QVKE---TT 125
QY 150 QPSGYTNIDKISNQDELNLNLPINYEYK-----ARPLSTTSQAQPSIKRVTVNQLAA 201
DB 126 QP-----QDNGSQRQVDLTTPKKVTQNGTETQVEAQPRTASEKPRVTR-SADVAEA 178
QY 202 EQGSNVNHLIKVTDQISITEGYDDSEGVIAK-----HDAENLIYDVTFEVDKVKSGD 253
DB 179 KEASDVSE-VKGTD--VTSKVTVESGIEAPQGNKVEPHAGQRVVLKPKADGLKRGD 235
QY 254 TMTVDIDKNTVPDLTDSFTIPKIDNS-----GEIATGTVDNKNKQITVFTDVKY 308
DB 236 YPFTLSNVNVTGVSTARKVPEIKGVSVMATGEILNG-----NRYTFNIEHK 288
QY 309 ENIKAHKLKTSYIDKSQVNN-----NTKLDVEYKTSALSSVNTKIIVFYORPNENRANL 363
DB 289 VEVANLEINLFDIPDKTVQSNGBQKITSKLNGE-----ETEKTIPTVYNPGVSNYTVN 342
QY 364 QSMFTNIDTKNHTVEQITLYINPLRYSAKETNVAISNGDSEGTIIDSTIIKVKY-VGDN 422
DB 343 NGSIETFNKESNKFTHIAVIKPMNGQSNST-VSVTGLTTEGSLAGQGPTRVYVYLKGGK 401
QY 423 QNLPSDNRIY----DYSEYEDVTDNDYDQAQGNNDNVNINFGNIDSPYIIKVISKYDPN 476
DB 402 DELPQS--VYANTSDTNKEKDVTKENNGKLSVQDNGSYSLNLDKDKTVIHYHTGEVLQ 459
QY 477 KDYTTIQOTVTMTQTTINET-----GEFRTASVDNTIAPSTSGGQGGDLPPEK 526
DB 460 SDQ-----VNPRTELYGYPERAYKSYVYVGGYR-LTWDNGGLVLYSNKADGNG----- 505
QY 527 TYKIGDYVWE---DVDKDGIONTNDNEKPLSNVLTLTYPDGTGTSKSV---RTDEGKYQ 579
```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:34:15 ; Search time 12.1685 Seconds
(without alignments)
2490.441 Million cell updates/sec

Title: US-10-806-288-13

Perfect score: 3013

Sequence: 1 SSDEKNDVNNQNSINTDD.....PDGTSKSVRTDEGKQFDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	300.5	10.0	1018	1 FNBA STAAU	P14738 staphylococ
2	203	6.7	537	1 ARP FLAFA	P04931 plasmodium
3	202	6.7	1093	1 P14K DCDI	P54677 dictyosteli
4	196.5	6.5	1139	1 HMW1 MYCGE	Q049413 mycoplasma
5	192.5	6.4	2867	1 RBP2 PLAVB	Q00799 plasmodium
6	192	6.4	918	1 YMJB CAEEL	P34487 caenorhabdi
7	189.5	6.3	1664	1 INT1 CANAL	P53705 candida alb
8	188.5	6.3	578	1 LIPA MYCPU	Q50274 mycoplasma
9	188.5	6.3	1435	1 EBAL PLAFD	P19214 plasmodium
10	186.5	6.2	3110	1 LMA2 HUMAN	P24043 homo sapien
11	184.5	6.1	1744	1 TANA XENLA	Q01550 xenopus lae
12	184	6.1	817	1 YG4A YEAST	P46949 saccharomyc
13	183.5	6.1	2009	1 SEC7 YEAST	P11075 saccharomyc
14	183	6.1	2452	1 RPB1 PLAFD	P14248 plasmodium
15	182.5	6.1	1096	1 MRCS YEAST	P25588 saccharomyc
16	182.5	6.1	1381	1 YBE7 YEAST	P34216 saccharomyc
17	182.5	6.1	1658	1 YM67 YEAST	Q03661 saccharomyc
18	182	6.0	1419	1 ALAL CANAL	O13368 candida alb
19	182	6.0	1460	1 N159 YEAST	P40477 saccharomyc
20	181.5	6.0	1093	1 YKDS CAEEL	Q03563 caenorhabdi
21	179.5	6.0	1165	1 YNF4 YEAST	P53950 saccharomyc
22	179	5.9	1253	1 DSPP HUMAN	Q9nz44 homo sapien
23	178	5.9	960	1 YMX6 YEAST	Q04279 saccharomyc
24	178	5.9	1875	1 MLPI YEAST	Q02455 saccharomyc
25	177	5.9	2660	1 YEEJ ECO57	Q8x8v7 escherichia
26	176	5.8	1196	1 BXCN CLOBO	P46081 clostridium
27	175.5	5.8	1183	1 CNA STAAU	Q53654 staphylococ
28	175.5	5.8	1637	1 MRSP STAAU	P80544 staphylococ
29	175	5.8	1147	1 CGAI HELPY	P80200 helicobacte
30	174	5.8	2022	1 ANTI1 ONCVA	P21249 onchocerca
31	173.5	5.8	967	1 Y5G0 CLOAB	P33747 clostridium
32	173	5.7	1310	1 YB35 SCHPO	Q14340 schizosacch
33	172	5.7	682	1 NISP LACLA	Q07596 lactococcus

ALIGNMENTS

RESULT 1

ID	FNBA STAAU	STANDARD;	PRT;	1018 AA.
AC	P14738;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Fibronectin-binding protein precursor (FNBP).			
GN	FNBA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NCIC 8325-4;			
RX	MEDLINE=89098998; PubMed=2521391;			
RA	Signaes C., Raucii G., Joensson K., Lindgren P.-E.,			
RA	Anantharamiah G.M., Hoeok M., Lindberg M.;			
RT	"Nucleotide sequence of the gene for a fibronectin-binding protein			
RT	from Staphylococcus aureus: use of this peptide sequence in the			
RT	synthesis of biologically active peptides.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).			
CC	- - FUNCTION: THE ABILITY OF FACTOR TO BIND FIBRONECTIN HAS BEEN			
CC	PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE			
CC	WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO			
CC	THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S.AUREUS,			
CC	THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE			
CC	INVASION.			
CC	- - SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; J04151; AAA26632.1; -			
DR	InterPro; IPR004237; Fn bind.			
DR	InterPro; IPR005877; Gpos YSIRK.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Ffam; PF02986; Fn_bind; 1.			
DR	Ffam; PF00746; Gram_pos_anchor; 1.			
DR	Ffam; PF04650; YSIRK_signal; 1.			
DR	TIGRFAMs; TIGR01167; LPXTG_anchor; 1.			
DR	TIGRFAMs; TIGR01168; YSIRK_signal; 1.			
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1. Signal.			
FW	Cell wall; Peptidoglycan-anchor; Repeat; 36			
KW	SIGNAL 1 36			
FT	CHAIN 37 985 FIBRONECTIN-BINDING PROTEIN.			
FT	PROPEP 986 1018 REMOVED BY SORTASE (POTENTIAL).			
FT	REPEAT 545 574 B-1.			
FT	REPEAT 575 604 B-2.			
FT	DOMAIN 745 878 4 X APPROXIMATE TANDEM REPEATS,			

Q02510 staphylococ
P53550 saccharomyc
P46590 candida alb
P07866 saccharomyc
P25558 saccharomyc
P47024 saccharomyc
P21131 streptococ
P25062 halobacteri
Q02192 streptococ
Q00798 plasmodium
P14164 saccharomyc
P47134 saccharomyc

34 171.5 5.7 688 1 LIP STAEP
35 170.5 5.7 970 1 PSUI_YEAST
36 170.5 5.7 1260 1 ALSI_CANAL
37 170.5 5.7 1435 1 LUB1_YEAST
38 170.5 5.7 1636 1 BUD3_YEAST
39 170 5.6 1803 1 YJL3_YEAST
40 169.5 5.6 571 1 TACY_STRPY
41 169.5 5.6 827 1 CSG_HALVO
42 169.5 5.6 1020 1 BCA_STRAG
43 169.5 5.6 2869 1 RBP1_PLAVB
44 169 5.6 731 1 BAF1_YEAST
45 169 5.6 954 1 BIR1_YEAST

```

FT REPEAT 745 782 FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 783 820 D-1.
FT REPEAT 821 859 D-2.
FT REPEAT 860 878 D-3.
FT DOMAIN 879 948 D-4 (INCOMPLETE).
FT REPEAT 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 879 948 WR1.
FT REPEAT 893 906 WR2.
FT REPEAT 907 920 WR3.
FT REPEAT 921 934 WR4.
FT REPEAT 935 948 WR5.
FT SITE 982 986 LPXG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 985 985 AMIBG-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;

Query Match 10.0%; Score 300.5; DB 1; Length 1018;
Best Local Similarity 21.2%; Pred. No. 1.1e-05;
Matches 123; Conservative 118; Mismatches 243; Indels 97; Gaps 25;

QY 35 GIEKRSDETSNTVNDENATFLQKTPQDN---THLTBEVKSSVSSSSIDTAQ 90
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 QDKEAAASEQKTTVEEN-----GNSATDNKTSQTATTNVNHIETQSYNAV--TE 82
QY 91 QPSHTT-INREE---SVQTSNVEDSHVSDPANSKIKESNTESKEENTIEOPNKYKEDS 146
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 QPSNATQTTBEAKVAQAPQAPQANIEVKEEVKE-----EAKPQVKETT 130
QY 147 TTSQPSG---YTNIDEKISNODELNLPINEYENKARPLSTSAQPSIKRVT-----V 196
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 QSQDSNGDQORQVLTPKKATQNAETQV---EVAQPTASESKPRTVSADVAEAKA 186
QY 197 NQLAAEGSNVNHILKVTQSGITEGDDSEGVKAHDAENLIYDVTFEVDKVKSGDTMT 256
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 SNAKVETGTDVTSKVTVEIGSI-EGHNNTKV-EPHAGQRAVLKYLKFLGHLHQGDYD 244
QY 257 VDIDKNTVPSDLTDSFTPIKONS-----GEIATGTDNKNKQIYTTDYDKYENI 311
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 FTLSNNVNTGVSRTARKVPELKGVSVMATGEVLEGG-----KIRYTFNDIEDKVDV 297
QY 312 KAHKLKLSYIDKSKVPNNNTKLDVEYKLTALSSV-----KTIIVYORPNEKNTANLQSM 366
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 TAELEINLFDIPKTVQTN-----QTITSTLNEEQTSKELDVKYKGIGNYANLGS 351
QY 367 FTNIDTQKHTVEQTIYINPLYSAKETNVAINSGDEGSIIDSTIIKVK-VGNQNL 425
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 IETFNKANNRFSHVAFIKP--NNGKTTSVTVTGLMKGSQNGQPKVRIFFELGNEDI 409
QY 426 PDSNRIY---DYSEYEDVTNDYAQLG--NNNDVNIENFGNIDSPYIIKVISYDPNKO- 478
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
410 AKS--VYANTTDSKFKVETSNMGNLNLQNGSYSNLNLENLQKTVVHYDGEYLGTD 467
QY 479 -DYTTIQOVTMOTTINEYTGFEPTASNDNTIAFSTSSGOGGDLPP-----EKTYKIGD 532
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
468 VDERT-QMVGHPEQLYKYDYDRGYTLTWIDNGLVLYSKANGNEKNGPIIQNKFEYK--- 523
QY 533 YVNEVDVKDGIQNTDNEKPLSNVLTLTYPDGTGSKSVRTD 573
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
524 ---EDTIKETLTQYDKN-----LVTTBEEYDSSTLDD 555

RESULT 2
ARP_PLAFA
ID ARP_PLAFA STANDARD; PRT; 537 AA.
AC P04931;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Asparagine-rich protein (AG319) (ARP) (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=86176787; PubMed=2421257;
RA Stahl H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,
RA Brown G.V., Anders R.F., Kemp D.J.;
RT "An asparagine-rich protein from blood stages of Plasmodium
RL falciparum shares determinants with sporozoites.";
CC Nucleic Acids Res. 14:3089-3102(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24328; AAA29491.1; -.
DR PIR; A23770; A23770.
DR InterPro; IPR008928; Glyco_trans_6hp.
KW Malaria.
FT NON_TER 1 1
FT TER 537 537
SQ SEQUENCE 537 AA; 63000 MW; A6D76B15318CC239 CRC64;

Query Match 6.7%; Score 203; DB 1; Length 537;
Best Local Similarity 17.8%; Pred. No. 0.054;
Matches 93; Conservative 106; Mismatches 212; Indels 112; Gaps 19;

QY 11 NNNQSTNDNNQIKKEETNNYDGEKSESDRTSTTNVDENATFLQKTPQDNTHL 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 NKNNNNDGNN--INYQNTNEFKNNKNNNNKNNNNKNNNNNNNNNNNNNNNNNNNN 61
QY 71 EYKSSSVSSSSSSID--TAQPSHTTT-----NRESVOT---SDNVEDSHVSDFAN 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 EHLR-NSIDMNSNINNYTNQTRFSSPMENENENENKNNYHGTGNNNNHFKYKVDNN 120
QY 120 SKIKESNTSGKEENTIEOPNKVSDSTTSQPSGYTNIDEKISNODELNL-PINEYENK 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 SSMK--NTDNKNTDTSYNNKGTINNDN-----NNMDYLRNINNNINEYKGS 163
QY 179 ASPLSTSAQPSIKKVTNNQALAAEQSNVNHILKVTQSGITEGDDSEGVKAHDAENLI 238
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 AKKKFTN-----YNNKNNLKFQNNNDNNNNNNNNNNNNNNNNNNNNNNNNNN 217
QY 239 YDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKONSGBIATGTVDNKNKQIT 298
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 RNNINIKENLN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 268
QY 299 YTFDYDVKYENIKAHKLKLSYIDKSKVPNNNTKLDVEYKLTALSSV----- 345
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
269 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 328
QY 346 -KTIIVYORPNEKNTANLQSM-----FTNIDTKNHTV-----EQTIYIN----- 384
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
329 VKMSMNVYNN-NESTANPQNNQNFQTNNDNNKNNNNNNNNNNNNNNNNNNNN 387
QY 385 -----FLYSAKETNVNI-----SGNGD-----EGSTIIDSTIIKVKYK-----VG 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
388 FSRAGYNNYNNLNNNNNNNSAVNNNSNGNNNNNNNNNNNNNNNNNNNNNNNN 447
QY 421 DNQNLPSDNRIYDYSEY-----EDVTNDDYAQLGNNDVNNIN 457
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
448 MNESINNNNTLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 490

RESULT 3
PI4K_DICDI
ID PI4K_DICDI STANDARD; PRT; 1093 AA.
AC P54677;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol 4-kinase (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-kinase) (PI4K-alpha).
```

GN PIKD OR PIK4.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takegawa K., Emr S.D., Firtel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
RT discoideum: biological roles of putative mammalian p110 and yeast
RT Vps34p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC -/- FUNCTION: Acts on phosphatidylinositol (PtdIns) in the first
CC committed step in the production of the second messenger
CC inositol-1,4,5-trisphosphate.
CC -/- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC 1-phosphatidyl-1D-myo-inositol 4-phosphate.
CC -/- SIMILARITY: Belongs to the PI3/P14-kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U23479; AAA85725.1; -;
DR PIR; T18275; T18275.
DR DictyBase; DDB0001907; pikD.
DR InterPro; IPR008938; ARM.
DR Pfam; PF00454; P13_P14_kinase; 1.
DR SMART; SM00146; PI3Kc; 1.
DR PROSITE; PS00915; P13_4_KINASE_1; 1.
DR PROSITE; PS00916; P13_4_KINASE_2; 1.
DR PROSITE; PSS0290; P13_4_KINASE_3; 1.
KW Transferase; Kinase; Multigene family.
FT DOMAIN 146 156 ASP-RICH.
FT DOMAIN 184 200 POLY-ASN.
FT DOMAIN 203 206 POLY-ASN.
FT DOMAIN 227 233 POLY-ASN.
FT DOMAIN 277 315 POLY-ASP.
FT DOMAIN 441 445 POLY-ASP.
FT DOMAIN 455 466 POLY-THR.
FT DOMAIN 494 501 POLY-GLY.
FT DOMAIN 690 696 POLY-THR.
FT DOMAIN 751 754 POLY-GLN.
FT DOMAIN 761 772 POLY-THR.
FT DOMAIN 775 785 POLY-THR.
FT DOMAIN 833 1093 PI3K/PI4K.
SQ SEQUENCE 1093 AA; 123017 MW; 8762BC78355AA635 CRC64;

Query Match 6.7%; Score 202; DB 1; Length 1093;
Best Local Similarity 20.0%; Pred. No. 0.14;
Matches 116; Conservative 100; Mismatches 223; Indels 142; Gaps 24;

QY 1 SSDEKNDVNNQSIINTD-DNQIITKEETNNYDGIKESRDETESTTVD-ENEATFL 58
DB 184 NSNNNNNNNNNSNDNNNNNNNNNNELIPNENSDNSINDENQYGNNSNNNNISGNNNIKI 243
QY 59 QKTPQDNTLTHEEVEKSSSVSSNSSI--DTAQCP-----SHTTINREESVQT 105
DB 244 DINSON-----KSDSIETLNSLTCEETKTSPIKODMNNNNNNNNNNNNNNNNNN 294
QY 106 SDNVEDSHVDFANSKIKESN-----TESGKEENTIQNKVKEDSTTSQPSG 153
DB 295 NNNNNNNNN--NNNNNNNNNNYGHINGSLSLDGIQGYISQPNDFIENTITQLKRN 351
QY 154 ---YTIIDEKISNQDELNLPINEYENKAPLSTTSQAQPSIKKVTYNQLAAGQSNVNL 210
DB 352 RIYKKVKEKELATRLRECFECSVITCSRPLITRPTSLPSPLISYNSGKIGGYHKI 411

QY 211 IKVTQOSTEGYDDSEGVKAHAENLIYDVTFFVDDVKVSGD-----TMTVDIDKNTVP 265
DB 412 LSPSSVDSTSLISEDDKIIEKEENVVED---DDDEVNSEDFIPTATTATTTTIP 468
QY 266 SLDLTSFTIPKIKDNGS---EIIATCTYDNKNKQIYTFYDVKYENIKAHKLKLSYID 322
DB 469 NHLKSTTSGVGINSSSTFININSAGAGGGEINHI-----GYDDI-----SYLD 515
QY 323 KSKVNNNTKL--DVEYKTA-----LSSVNTKITVE---YORPNNERTANLOS 365
DB 516 KCKTPPAESKSDHDFEPFLSKSHRCDYLDNDILSPFQKLAHISKILLPIPIDLRQAKLKH 575
QY 366 MFTNIDTQHTVEQIYINPLRYSAKETNNVNSGNGDEGSTIIDDSTIIKYKVGDNQNL 425
DB 576 ---EISLLNINPLGLYV-PLWQSS-----NHHCVVRI-----P 605
QY 426 PDSNRIYDSE-----YEDVTNDVYAQLGNNDVNNFGNIDSPYIIKVISKYDPNKDD 479
DB 606 PREVKILNSRRVFPFLVLEVIESEHEALSSN-----IFEVVSY-----LQ 647
QY 480 YTTIOQTVMQTTINEYTG-----FRASYDNTIAFTSS 515
DB 648 YTTGNSALKKDDIKRKYSEKFKKSLNSSINSTISNDS 688

RESULT 4
HMW1 MYCGE STANDARD; PRT; 1139 AA.
ID AC Q494I3; Q49365;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytoadherence high molecular weight protein 1 (Cytoadherence accessory
DE protein 1).
GN HMW1 OR MG312.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 721-847 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075330; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bost K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing."
RL J. Bacteriol. 175:7918-7930(1993).
CC -/- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW
CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHESIN PROTEINS
CC IN THE MYCOPLASMA MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
CC SIMILARITY).
CC -/- SUBCELLULAR LOCATION: Localizes specifically to the attachment
CC membrane (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

QY 182 LSTTSAPQSKRVTV--VNLQAAE-QGSN-----VNHILKVTQOSTEGVDDSEGVIKAHDAE 235
 Db 592 NSSNSSTNS-KEFTDKLKELETBEGLNKTGKYLQIBEDIKVKNEDRS---LKNQIEQ 647
 QY 236 NLIYDVTFFVDDVKVS---GDMTMTVDIDK---NTVPSDLTDSPTPIKIKNSGEIIA 286
 Db 648 HLKY--TSDNRDNNKTLISKNDEIQKIEKILNDAPSG-KDKFTTEK----- 695
 QY 287 TGYDNNKQITVFTFDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVYKT--ALSSV 344
 Db 696 --NLQNKVKI-----IDEPKEDLQLLNSL---SKYEHEQKLYNEASTIEKIKOL 743
 QY 345 NKITVEYQRPENRTAN-----LQSMFTNIDT-----KNTVZOTYINPLRYSAKETNV 395
 Db 744 HQTKEEYEKLEKMKFNSFGQILDKLNTLDELNLTLEKNIVEQTNVINKV---MSDSL 800
 QY 396 NISNGDEGSTIID---DSTILKVKVGDNQ-----NLPDSNRIY-DYSE 436
 Db 801 NLTAEDNLRSALDGYRADSTELTKYKNRINERKEKELSTLKQEDDIPGKNIYEYNN 860
 QY 437 YEDV-TNDDYAOQLGNNDVNFNGNIDSPYIIKVISKYDPNKDDYTTIQOTVTMOT---- 491
 Db 861 HKNVMNKEHKISSDINQCNEN-----IIKAENLSTFNTLVQTLDAHTGKCD 908
 QY 492 -TNEYTGEPRT 502
 Db 909 QKVHDLQKEFT 920

RESULT 6
 YMJCB CAEEL
 ID YMJCB CAEEL STANDARD; PRT; 918 AA.
 AC P3487;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein F59B2.12 in chromosome III.
 GN F59B2.12.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer B., Staden K.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

DR EMBL: Z11505; CAA77581.1; --
 DR PIR: G88545; G88545.
 DR WormPep; F59B2.12; CE01024.

KW Hypothetical protein.

SQ SEQUENCE 918 AA; E464FD86B14945DE CRC64;
 Query Match 6.4%; Score 192; DB 1; Length 918;
 Best Local Similarity 21.2%; Pred. No. 0.29;
 Matches 138; Conservative 90; Mismatches 278; Indels 146; Gaps 26;
 QY 4 EKKDVINNQSIINTDDNNQIIKKEETNYD-----GIEKSESDRTE-----STTN 49
 Db 113 ESKVDEANENTIKISADGS-VIETGKSHNKSDDASSYGLEKSKTYADKNGTMLSSNTN 171
 QY 50 VDNEEA-----TFLOKTPQDNTHL-----TEEEVKSSSVSESSNSIDTAQOP 92
 Db 172 KINQGRSAALDEGNEFVQGNADGTFLRNNTGCHKNTDEHLSNVLDENAOQSIG-ADGT 230
 QY 93 SHTTINREBSVQTSN-VEDSHVSDPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQP 151
 Db 231 SHNITNRKSGVGDHNAASDAH-----SNFESLDAQGNKKSQNSYKKAASG- 278
 QY 152 SGVTNIDEKISNODELLNLPINEYENKARPLSTTS-----AQPSIKRVTN-----QLA 200
 Db 279 ---SNADFE-SNLESILKNADGTSMSNSTGNFNNTSYDKATAEEVMSKKNVADGTSMEA 334
 QY 201 AEOGSSNVNHLIKVTDQSIITEGYDDSEGVKAHDAENLIYDVTPEVDVKVSGDTMTVDID 260
 Db 335 SHAGSSNKSINSASGSSDLSWVGPNQ-ISKHSTSNKTDN--YALDEANQSGSISQIG 391
 QY 261 K-----NTVPSDLTDSPTPIKIKNSGEIIATGYDNNKKNQIYTT 300
 Db 392 KNGQSLNESSIESGRKAESRNNNTAATLDSVDANGTVSSSHSKSASGTSLDENHNKTHA 451
 QY 301 PTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVYKTA---LSSV-----KTIIV 350
 Db 452 LQASVDEHGNMKNHSDIGSYRNKKKTGEPGNSASASIKKADGTMSSQNVNKNNTNRNTYEA 511
 QY 351 EYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNLSNGDSEGSTIIDD 410
 Db 512 EKSALEKNHEKNSDGTFKD-----ESKGSNSRVN-RTDGSNLAUG 551
 QY 411 STIIKVKYKVDNQNLPSNRIDYSEYEDVTNDDYAOQLGNNDVNFNGNIDSPYIKVI 470
 Db 552 SYSVGVGGVSSNETIASSN-AFNWSDAES-NQFDHLHQKTANGTEITHAK-DS---KQV 604
 QY 471 SKYDPNKKDYTTIQOTVTMOTTINEYTGFRTRASYDNTAFSTSSGCGQGLDPEKTYKI 530
 Db 605 AASANAKSLDTSMSAVDAKGNKVDKTSQAADSHD---AISASS----- 646
 QY 531 GDYVDEVDKDGQIQTNDNNEKPLSNVLVTLTYPDGTSKSVRTDEDDGKYQPDG 582
 Db 647 -----DVEDAKIVKHA-DRSEISNDSNQTAHEHNDSSKQSEHKRQNDG 691

RESULT 7

INTL CANAL STANDARD; PRT; 1664 AA.

ID INTL CANAL
 AC P53705;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Integrin alpha chain-like protein (Alpha-INT1).
 GN INT1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10261;
 RX MEDLINE=96133936; PubMed=8552638;
 RA Gale C., Finkel D., Tao N., Meinke M., McClellan M., Olson J.,
 RA Kendrick K., Hostetter M.;
 RT "Cloning and expression of a gene encoding an integrin-like protein
 RT in Candida albicans.";

Proc. Natl. Acad. Sci. U.S.A. 93:357-361(1996).

-!- FUNCTION: Could play a role in adhesion and in STE12-independent morphogenesis.

-!- SUBCELLULAR LOCATION: CELL-SURFACE OF THE BLASTOSPORES.

-!- SIMILARITY: SOME, WITH INTEGRIN ALPHA CHAINS.

-!- SIMILARITY: Contains 1 PH domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U35070; AAA96019.1; .

PIR; T18216; T18216.

InterPro; IPR001849; PH.

Pfam; PF00169; PH; 1.

SMART; SM00233; PH; 1.

PROSITE; PS50003; PH DOMAIN; 1.

Cell adhesion; Glycoprotein.

DOMAIN 1527 1636

FT SITE 1149 1151

FT CARBOHYD 50 50

FT CARBOHYD 154 154

FT CARBOHYD 216 216

FT CARBOHYD 223 223

FT CARBOHYD 265 265

FT CARBOHYD 268 268

FT CARBOHYD 285 285

FT CARBOHYD 289 289

FT CARBOHYD 306 306

FT CARBOHYD 314 314

FT CARBOHYD 446 446

FT CARBOHYD 583 583

FT CARBOHYD 638 638

FT CARBOHYD 670 670

FT CARBOHYD 691 691

FT CARBOHYD 817 817

FT CARBOHYD 838 838

FT CARBOHYD 841 841

FT CARBOHYD 1018 1018

FT CARBOHYD 1082 1082

FT CARBOHYD 1100 1100

FT CARBOHYD 1103 1103

FT CARBOHYD 1113 1113

FT CARBOHYD 1200 1200

FT CARBOHYD 1571 1571

FT CARBOHYD 1593 1593

FT DOMAIN 95 106

FT DOMAIN 283 286

FT DOMAIN 1283 1292

FT DOMAIN 1651 1660

SQ SEQUENCE 1664 AA; BA2EFDDC8196790 CRC64;

Query Match 6.38; Score 189.5; DB 1; Length 1664;

Best Local Similarity 20.44; Pred. No. 0.75;

Matches 131; Conservative 105; Mismatches 236; Indels 171; Gaps 31;

QY 8 DVINNQS---INTDDNNQIIKK-----EETNNYDGIEKRSBDRSTET 48

DB 138 DKNENHAPYINTSPNKSIMKATPKASPCKVAVFTVTPNPEIHYP--DNVVEEEDSQ 195

QY 49 NVDENATFLQKTPQDNTHTEEEVKSSSVESNSIDTAQPSHTT-----INREE 101

DB 196 KEDSVPEPLIQ-----HQMCKQSFNYSDNEDTASVPPT-PPLTKPTFAQLNKN 247

QY 102 SVQTSNDVEDSHVSD-----FANSKIKE-----SNTESGKEHTIEQPKVKEDST 148

DB 248 EV-----NSEPEALTMKLENFSLDEKNVLYLSPNTNNNSKNVSDMSHLQ-----N 299

QY 149 SQPSGYNTIDEKISNQDELNLNLPINEYENKARPL-STTSAQPSIKRVTYNQLAAEQGSNV 207

Db 300 LQDASKNKTENIHNLSPALKAPKNDIEN---PLNSITNADISLRSSGSSQSLQSLRND 356

QY 208 NHLIKV---TDQSITEGYDDSEGVIAKHAENLIYDVTTEVDKVKSGDTMTVDIDKNTV 264

Db 357 NRVLSEVPGSPKKVNFGLNDGKGFSD-----EVVESLL 392

QY 265 PSDLT-DSFTPIKID-----NSGEIIATGYDNKNKQIYTFDYVDKYENIKAH----- 314

Db 393 PRLSRDKLETTKEHAPEHNENFIDAKSTNTNKG-QLLVSSDDHDSFDRSYNHTEQS 451

QY 315 -LKLTSYIDKSKVPNN-----NTKLDVEYKTALSSVVK 346

Db 452 ILNLLNSASQISLNALEKQRTQEQEQTQAAPEEETSFSdniVKVQPKSNLEFVKV 511

QY 347 TITVEYQRPNEFTANLQSMFTNIDTNKH-TVEQTIYINPLYSAKETVNNISG----- 399

Db 512 TIKKEPVSAATEIKAPKREFSSRLIRIKNEDETAEPADIIHPKKEANEANSHVEDTALLKKA 571

QY 400 -NGDEGSTIIDSTIIKV-YKVGDNQNLPSNRIYDYSEVEDVTDNDYLAQLGNN---NDV 454

Db 572 LNDDESDTTONSTKMSIRFHIDSDWKLESDND---GDRED--NDDISRFKSDILNDV 625

QY 455 N-----INFGNIDSPYIKVISKYDPNKKDYTTTIOQTVMQTTINEYTGFR-----T 502

Db 626 SQTSDIIGDKYGNSSSEITTKTLA---PPRSNDNDKNSKSLSDPANNESLQQOLEVPHT 682

QY 503 ASVDNTIAFSTSGGQGGDLPRPEK---TYKIGDY-VWEDVDK 540

Db 683 KEDDSILANSSNIA-----PPEELTLPVVEANDYSSFNNDVTK 719

RESULT 8

LIPA MYCPU STANDARD; PRT; 578 AA.

AC Q50274; Q90Q39;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Lipoprotein A precursor.

GN LIPA OR MYPU 5300.

OS Mycoplasma pulmonis.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2107;

RN [1] SEQUENCE FROM N.A.

RC STRAIN=KD735-15;

RX MEDLINE=20245550; PubMed=10781561;

RA Shen X., Gumalak J., Yu H., French C.T., Zou N., Dybvig K.;

RT "Gene rearrangements in the vsa locus of Mycoplasma pulmonis.";

RL J. Bacteriol. 182:2900-2908(2000).

RN [2] SEQUENCE FROM N.A.

RC STRAIN=UAB CTIP;

RX MEDLINE=21267165; PubMed=11353084;

RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galissou F.,

RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,

RA Blanchard A.;

RT "The complete genome sequence of the murine respiratory pathogen

RT Mycoplasma pulmonis.";

RL Nucleic Acids Res. 29:2145-2153(2001).

RN [3] PARTIAL SEQUENCE FROM N.A.

RC STRAIN=KD735-15;

RX MEDLINE=96414471; PubMed=8817492;

RA Bhugra B., Voelker L.L., Zou N., Yu H., Dybvig K.;

RT "Mechanism of antigenic variation in Mycoplasma pulmonis: interwoven,

RT site-specific DNA inversions.";

RL Mol. Microbiol. 18:703-714(1995).

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor

CC (Potential).

CC -!- SIMILARITY: BELONGS TO THE PULMONIS LIPAB LIPOPROTEIN FAMILY.

QY 391 K---ETNVNIGSGDEGTTIIDSTIIKVKYKVGDNQNLPSNRRIYDYSEYEDVTNDYQAQ 447
 Db 1235 KHGFHTMNLHGCG-----VSRQSHNSHHGNRQD-----R 1266
 QY 448 LGNNNDV-----NINFGNIDSPYIKVISKVD-----PNKDDYTTIOQTVMQTTI 493
 Db 1267 GGNISGNVLNMGKSNNNFNIPSRV-----NLYDKKLDLDLYENRNDSTT-KELIKKLAEI 1320
 QY 494 NEYTGEFRTASVDNTI 509
 Db 1321 NKCENISVKYCDHMI 1336

RESULT 10
 ID LMA2 HUMAN
 AC P24043; Q14736; STANDARD; PRT; 3110 AA.
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Laminin alpha-2 chain precursor (Laminin M chain) (Merossin heavy chain).
 DN LMA2 OR LAM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94124633; PubMed=8294519;
 RA Vuolteenaho R., Niesinen M., Sainio K., Byers M., Eddy R., Hirvonen H., Shows T.B., Sariola H., Engvall E., Tryggvason K.; "Human laminin M chain (merossin): complete primary structure, chromosomal assignment, and expression of the M and A chain in human fetal tissues." J. Biol. Chem. 271:27664-27669(1996).
 RL J Cell Biol. 124:381-394 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97066955; PubMed=8910357;
 RA Zhang X., Vuolteenaho R., Tryggvason K.; "Structure of the human laminin alpha2-chain gene (LMA2), which is affected in congenital muscular dystrophy." J. Biol. Chem. 271:27664-27669(1996).
 RL J. Biol. Chem. 271:27664-27669(1996).
 RN [3]
 RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=90238994; PubMed=2185464;
 RA Ehrig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.; "Merossin, a tissue-specific basement membrane protein, is a laminin-like protein." Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268(1990).
 RL [4]
 RP VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.
 RA Panicker S.G., Mendell J.T., Chen L., Peng B., Sahenk Z., Marzluf G.A., Anato A.A., Mendell J.R.; "Novel single base polymorphisms and rare sequence variants in the laminin 2-chain coding region detected by RNA/SSCP analysis." Hum. Mutat. 13:174-174(1999).
 RL [5]
 RP ERRATUM.
 RA Panicker S.G., Mendell J.T., Chen L., Peng B., Sahenk Z., Marzluf G.A., Anato A.A., Mendell J.R.; Hum. Mutat. 13:340-340(1999).
 RL [6]
 RP VARIANT MDC1A PRO-2564.
 RX MEDLINE=21476011; PubMed=11591858;
 RA He Y., Jones K.J., Vignier N., Morgan G., Chevallay M., Barois A., Estournet-Mathaud B., Hori H., Mizuta T., Tome F.M.S., North K.N., Guicheney P.; "Congenital muscular dystrophy with primary partial laminin alpha-2 chain deficiency: molecular study." Neurology 57:1319-1322(2001).
 [7]
 RP VARIANTS MDC1A TYR-527 AND ARG-862.
 RX MEDLINE=22439669; PubMed=12552556;
 RA Tezak Z., Prandini P., Boscaro M., Marin A., Devaney J., Marino M., Fanin M., Trevisan C.P., Park J., Tyson W., Finkel R., Garcia C., Angelini C., Hoffman E.P., Pegoraro E.; "Clinical and molecular study in congenital muscular dystrophy with partial laminin alpha-2 (LMA2) deficiency." Hum. Mutat. 21:103-111(2003).
 RL [8]
 CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin. is thought to mediate the attachment, migration and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components.
 CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three different polypeptide chains (alpha, beta, gamma), which are bound to each other by disulfide bonds into a cross-shaped molecule comprising one long and three short arms with globules at each end. The alpha-2 chain is a subunit of laminin-2 (merossin) and laminin-4 (S-merossin).
 CC -!- SUBCELLULAR LOCATION: Extracellular; found in the basement membranes (major component).
 CC -!- TISSUE SPECIFICITY: Placenta, striated muscle, peripheral nerve, cardiac muscle, pancreas, lung, spleen, kidney, adrenal gland, skin, testis, meninges, choroid plexus, and some other regions of the brain; not in liver, thymus and bone.
 CC -!- DOMAIN: The alpha-helical domains I and II are thought to interact with other laminin chains to form a coiled coil structure.
 CC -!- DOMAIN: Domains VI, IV and G are globular.
 CC -!- DISEASE: Defects in LMA2 are the cause of merossin-deficient congenital muscular dystrophy type 1A (MDC1A) [MIM:607855]. MDC1A is characterized by difficulty walking, hypotonia, proximal weakness, hyporeflexia, and white matter hypodensity on MRI.
 CC -!- SIMILARITY: Contains 1 laminin N-terminal domain.
 CC -!- SIMILARITY: Contains 17 laminin EGF-like domains.
 CC -!- SIMILARITY: Contains 5 laminin IV domains.
 CC -!- SIMILARITY: Contains 5 laminin G-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL; Z26653; CAA81394.1; -
 DR EMBL; U66796; AAB18388.1; -
 DR EMBL; U66733; AAB18388.1; JOINED.
 DR EMBL; U66734; AAB18388.1; JOINED.
 DR EMBL; U66735; AAB18388.1; JOINED.
 DR EMBL; U66736; AAB18388.1; JOINED.
 DR EMBL; U66737; AAB18388.1; JOINED.
 DR EMBL; U66738; AAB18388.1; JOINED.
 DR EMBL; U66739; AAB18388.1; JOINED.
 DR EMBL; U66740; AAB18388.1; JOINED.
 DR EMBL; U66741; AAB18388.1; JOINED.
 DR EMBL; U66742; AAB18388.1; JOINED.
 DR EMBL; U66743; AAB18388.1; JOINED.
 DR EMBL; U66745; AAB18388.1; JOINED.
 DR EMBL; U66746; AAB18388.1; JOINED.
 DR EMBL; U66747; AAB18388.1; JOINED.
 DR EMBL; U66748; AAB18388.1; JOINED.
 DR EMBL; U66749; AAB18388.1; JOINED.
 DR EMBL; U66750; AAB18388.1; JOINED.
 DR EMBL; U66751; AAB18388.1; JOINED.
 DR EMBL; U66752; AAB18388.1; JOINED.
 DR EMBL; U66753; AAB18388.1; JOINED.
 DR EMBL; U66754; AAB18388.1; JOINED.
 DR EMBL; U66755; AAB18388.1; JOINED.
 DR EMBL; U66756; AAB18388.1; JOINED.
 DR EMBL; U66757; AAB18388.1; JOINED.
 DR EMBL; U66758; AAB18388.1; JOINED.

DR ENBL; U66759; AAB18388.1; JOINED.
DR ENBL; U66760; AAB18388.1; JOINED.
DR ENBL; U66761; AAB18388.1; JOINED.
DR ENBL; U66762; AAB18388.1; JOINED.
DR ENBL; U66763; AAB18388.1; JOINED.
DR ENBL; U66764; AAB18388.1; JOINED.
DR ENBL; U66765; AAB18388.1; JOINED.
DR ENBL; U66766; AAB18388.1; JOINED.
DR ENBL; U66767; AAB18388.1; JOINED.
DR ENBL; U66768; AAB18388.1; JOINED.
DR ENBL; U66769; AAB18388.1; JOINED.
DR ENBL; U66770; AAB18388.1; JOINED.
DR ENBL; U66771; AAB18388.1; JOINED.
DR ENBL; U66772; AAB18388.1; JOINED.
DR ENBL; U66773; AAB18388.1; JOINED.
DR ENBL; U66774; AAB18388.1; JOINED.
DR ENBL; U66775; AAB18388.1; JOINED.
DR ENBL; U66776; AAB18388.1; JOINED.
DR ENBL; U66777; AAB18388.1; JOINED.
DR ENBL; U66778; AAB18388.1; JOINED.
DR ENBL; U66779; AAB18388.1; JOINED.
DR ENBL; U66780; AAB18388.1; JOINED.
DR ENBL; U66781; AAB18388.1; JOINED.
DR ENBL; U66782; AAB18388.1; JOINED.
DR ENBL; U66783; AAB18388.1; JOINED.
DR ENBL; U66784; AAB18388.1; JOINED.
DR ENBL; U66785; AAB18388.1; JOINED.
DR ENBL; U66786; AAB18388.1; JOINED.
DR ENBL; U66787; AAB18388.1; JOINED.
DR ENBL; U66788; AAB18388.1; JOINED.
DR ENBL; U66789; AAB18388.1; JOINED.
DR ENBL; U66790; AAB18388.1; JOINED.
DR ENBL; U66791; AAB18388.1; JOINED.
DR ENBL; U66792; AAB18388.1; JOINED.
DR ENBL; U66793; AAB18388.1; JOINED.
DR ENBL; U66794; AAB18388.1; JOINED.
DR ENBL; U66795; AAB18388.1; JOINED.
DR ENBL; MS9832; AAA63215.1; -.
DR PIR; PX0082; MGHUMH.
DR HSSP; Q60675; 1QUO.
DR Genew; HGNC:6482; LAMA2.
DR MIM; 156225; -.
DR MIM; 607855; -.
DR GO; GO:0005604; C:basement membrane; TAS.
DR GO; GO:0005198; F:structural molecule activity; TAS.
DR GO; GO:0007517; P:muscle development; TAS.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR008979; EGF bind like.
DR InterPro; IPR000034; Laminin B.
DR InterPro; IPR002049; Laminin EGF.
DR InterPro; IPR001791; Laminin G.
DR InterPro; IPR008211; LamNT.
DR Pfam; PF00052; laminin B; 2.
DR Pfam; PF00053; laminin EGF; 14.
DR Pfam; PF00054; laminin G; 5.
DR Pfam; PF00055; laminin Nterm; 1.
DR PRINTS; PR00011; EGF/LAMININ.
DR ProDom; PD003031; Laminin B; 1.
DR SMART; SM00180; EGF Lam; 15.
DR SMART; SM00281; LamB; 2.
DR SMART; SM00282; LamG; 5.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01248; LAMININ TYPE EGF; 14.
DR PROSITE; PS50025; LAM G DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism;
KW Disease mutation.
FT SIGNAL 1 22 POTENTIAL.

Query Match 6.2%; Score 186.5; DB 1; Length 3110;
Best Local Similarity 19.9%; Pred. No. 2.1;

Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;
QY 21 NNQIIKKEET-----NNYDGIKESDRTSTTNVDENATFLOKTPQDNTLHTE 71
DB NEKAIKLNLTGLTRDBAFERNLEGLQK-----EIDQMIKELRRKNLETKQIEAD 1742
QY 72 EYKESSESSNSSIDTAQOPSHHTIN---RESVQTSNDVEDS-HVSPFANSKIKESN- 126
DB ELVAABALLKVKLFGESRGENEEMKDLREKLDYKNKVDDAWDLLREATDKREAR 1802
QY 127 -----TESGKE--ENTIEOPNKV-----KEDSTTS 149
DB LPAVQKQWTALEKKKEAVESGKQIENLTKEGNDILDEANLADAINSIIYVEDIQIK 1862
QY 150 QPSGYTNIDEKISN-----QDELLNLPINEYKNKAPLSTTSA-----QPSIKRVTVNQL 199
DB LPPMSEELNIDDLQSIEIKDRKLAEKVQSAESHAQLNDSAVLDGILDEAKNISFNAT 1922
QY 200 AAEQG-SNVNHLIKVTDQSITEGYDSEGVKAHAENLIYDVTPEVDDKVKSGDTMTVD 258
DB AAFKAYSNIKQYI-----DEAEKVAK--EAKDLAHEAT-----KLATGPR---G 1961
QY 259 IDKNTVPSDLTDSFTI-----PKIKNSGEIATGT-----YONKNKQITVTFDYVD 306
DB LKEDAKGCLQKSPRILNEAKKLANDVKENEDHLNGLKTRIEADARNGDLLRTLNDTLG 2021
QY 307 KYENI--KAHLKLTSYDKSKVNNNTK-----LQVEYKLTALSSVNNKTIT 349
DB KLSAIPNDTAAKLVQAKQKARQANDTAKDLVLAQITELHQNLGDKVKNKLADSVAKTNA 2081
QY 350 VEYQRPENR--TANLQSMETNID-----TKNHTVEQTIYINPLRYS 389
DB V-KDPSKKNKIADADATVKNLEQBADRLIDKLPKELEDNLKNKISEKELINOARKQ 2140
QY 390 AKETNVNISGNGD-----EGS-----TIIDSTIIKV-----YK 418
DB ANSIKVSVSSGGDCIRTYKPEIKKGSYNNIVNVKTAVADNLLFYLSAKFIDFLAEMR 2200
QY 419 VGDNQNLPSNRIYDYSEYEDVTND-----YALQGNNDVNINFGNIDSPYIIKVIS 471
DB KGKVSFLMDVSGVGVRVEYDPTIDDSYWRIVASRTGRNGTISVR--ALDGPKASIVS 2258
QY 472 -KYDPNKDDYTTIQQTVMTQTTINEYGEFRFASVDNTIAFTSSGQGGDLPPEKTYKI 530
DB THSTSPPGYTTILDVDANAMLFVGLTGKAKADAVRVITFTGCMGETYFDNKP-----I 2313
QY 531 GDYVWEDVDKGIQNTNDNEKPLSNVLVLTYPD--GTSKSVRT-DEGKYQFDG 582
DB G--LWNFREKEG-----DCKGCTVSPQVEDSEGTIQFDG 2345

RESULT 11
TANA XENLA
ID TANA XENLA STANDARD; PRT; 1744 AA.
AC Q01550;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanabin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tadpole head;
RX MEDLINE=92398961; PubMed=1524825;
RA Hemmati-Brivanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate
RT neurons defines a new class of intermediate filament protein.";
RL Neuron 9:417-428(1992).
CC -I- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.

CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURULA AND PERSIST
CC DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
CC CORD.
CC
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M99387; AAA49966.1; -
DR PIR; JH0720; JH0720.
DR InterPro; IPR001664; IP.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone.
FT DOMAIN 1 12 HEAD.
FT DOMAIN 13 314 ROD.
FT DOMAIN 315 1744 TAIL.
FT DOMAIN 8 48 COIL 1A.
FT DOMAIN 49 60 LINKER 1.
FT DOMAIN 61 156 COIL 1B.
FT DOMAIN 157 179 LINKER 12.
FT DOMAIN 180 193 COIL 2A.
FT DOMAIN 194 199 LINKER 2.
FT DOMAIN 200 314 COIL 2B.
SQ SEQUENCE 1744 AA; 199561 MW; 6502EAC9FE6C4E93 CRC64;

Query Match 6.1%; Score 184.5; DB 1; Length 1744;
Best Local Similarity 19.7%; Pred. No. 1.3;
Matches 149; Conservative 122; Mismatches 239; Indels 245; Gaps 42;

QY 1 SSDEKNDVNNOSINTDNNQIIKKEETNN--YDIEKRSDETRSTTNDENEATFL 58
DB 886 SEIEKDFINEQECLSKDQIRAPDTEEDVHQVDFPMQSFEREVGQNNIKQEVDFL 945
QY 59 QKTPQD-----NHLTEEEVK--ESSSV-----ESSNSSIDT- 88
DB 946 QNYDEDSFQNDPEQELSCDLEQKIKLEENQNSNEGQNFNGNDIEFQSGQGYD 1005
QY 89 -----AQPSHTTINREE-----SVQSDNVEDSHVDFANSKIKESN 126
DB 1006 EICQETIGNVSAQLLCESINDQDKLSEMEDEEQQNPETEDNIGLEQSDQENTRSNEG 1065
QY 127 TESGKENTIEQPKVKEDSTTQSPGYTNIIDEKISNQ--DELLNLPINE--YENKARPL 192
DB 1066 KFSQEECDVVFKP-----ED--MSDKSEYSGQEDLDKQVTFSLNQANNLLKEDEVIL 1119
QY 183 STTSAQPSI-----KRVTVNQLAAEQGSNNVNHILKVTQDSI--TEGY----- 222
DB 1120 HHADDQRSVNDIEITDEKLSRIIDNELAT--VDVNESLAANKQVOLFTEYAVDDNV 1176
QY 223 ---DDSEGVTKAH-----DAENLIYVTFE-----VDDKVKSGDTMTVDI 259
DB 1177 GMQDDSGGYQTKEDLPVDFGNNIIEKIEIQOTSLNQEICERVDNVDEDI--SGEAKNESV 1235
QY 260 DKNTVP-----SDLTDS-----FTIPKIKNSGEIATCTYDNKNKQIYTF 301
DB 1236 EMDVDVLVPEAKVTGDEQISPLQDEKLNLETMDTKNDGQQL-----CLKENE----- 1285
QY 302 TDYVDKYENIK-----AHLKLTYSIDKSKVP-----NNNTKL-----DVEYKLTALSS-- 343
DB 1286 TEYIEVTDSPQATDLSDHAGRELT--VDQNSANLQPCENPTKTLIAHIEYETVADSDL 1343
QY 344 -----VNKITVEYQRPENRNTANLQSMFTNTDKNHTVEQIYINPLYSKAKETVNI 397
DB 1344 ESTEEQVQETERIPP--KPEDSKMEN-----ENSEEESVDSQBSILN-----SHKSEFEI 1393
QY 398 SGNDEGSTIIDSTIIKVKVGDNQNLDPDSNRIYDY--SEYEDVTN--DDYAQLGNNDV 454

DB 1394 S-----KDYQL--EQTLFDVTPLPNLEDEFEFEDLTQPDVHEHQNDSDS 1435
QY 455 NINFGNIDSPYIIKISKYDPNKKDDYTTIQQTVMQTINEXT-GEFRTASYDNTIAFST 513
DB 1436 GA-----STFTSV-----DEKEREVRSVSKDEESNEEFQDVLSDKTSQVEVTT 1483
QY 514 SSGQGGDLPPEKTYKIGYVWEDVDKGIQNTND-NKPLSNVL-----557
DB 1484 LSGLAQ-----EFSY-LGD---NEESEDSENAEILNENPSNDIVDFMVMSQMTETKIIIA 1534
QY 558 -----VTLTYPDGTSK---SVRTDEDKGYQDF 581
DB 1535 EQVTEQTEVTLQFDDAPNKLITENLNAREKETDYIE 1569

RESULT 12
YG4A YEAST
ID YG4A YEAST STANDARD; PRT; 817 AA.
AC P46949;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 90.8 kDa protein in RRP41-SNG1 intergenic region.
GN YGR196C OR G7589.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96076633; PubMed=7502584;
RA Guerreiro P., Maia e Silva A., Barreiros T., Arroyo J.,
RA Garcia-Gonzalez M., Garcia-Saez M.I., Rodrigues-Pousada C.,
RA Nombela C.
RT "The complete sequence of a 9000 bp fragment of the right arm of
RT Saccharomyces cerevisiae chromosome VII contains four previously
RT unknown open reading frames.";
RL Yeast 11:1087-1091(1995).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DB EMBL; X82775; CAAS8019.1; -
DR EMBL; Z72981; CAA97222.1; -
DR PIR; S53919; S53919.
DR GerOnline; 141508; -
DR SGD; S0003428; FYV8.
DR GO; GO:0005737; C:cytoplasm; IDA.
KW Hypothetical protein.
SQ SEQUENCE 817 AA; 90797 MW; E52C5D659D63BEBB CRC64;

Query Match 6.1%; Score 184; DB 1; Length 817;
Best Local Similarity 21.0%; Pred. No. 0.54;
Matches 134; Conservative 88; Mismatches 237; Indels 180; Gaps 31;

QY 1 SSDEKNDVNN-----NQSINTDNNQI--IKKEETNNYDIEKRSDE 42
DB 120 SSEHKSQDYLSTASLKSPSENKSPHTNRAVNDLNLIEQISREMTPE---IROTSDF 176
QY 43 RSTSTNVD--NEATFLQKTPQDNTLHTEEVKSSSVSESSNSIDTAQPSHTTINRE 100
DB 177 RDS--DSCDEIQNEAPLGEAVFSSPVEDEKSHLGVSMNTNEADTF--NTFRNGN 233
QY 101 ESVQTSNDNVEDSHVDFANSKIKESN-----TESGKENTIEQPKVKEDSTTSPSGYT 155
DB 234 EHLSSDGDVSEQKDDDEF---KVSERGLADILPAKEEN-LQQ-----EDDGEVSSGAL 284


```

376 EEDSSSLCHGASLKAIRTIYVNFVFSNPSNOGIAQTLPQIISVVYDKIDLKQS----- 431
358 NRTANLQSMFTNIDTKNHTVEQTIIYINPLYSAKETNNISGNGDEGSTIIDDSTIIKVV 417
432 -----TSSAVSLTKNKHQQAELS-----EASENETAPL-----TLENMD 470
418 KVGONQNPDSNRIVDYSEYED--VTNDYYA 446
471 KLND-----DEERLMDAQQPSIAITNQDLA 496

RESULT 14
ID RPB1_PLAFD STANDARD; PRT; 2452 AA.
AC P14248;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6).
GN RPII.
OS Plasmodium falciparum (isolate CDC / Honduras).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90098832; PubMed=2690004;
RA Li W.B., Bzik D.J., Gu H., Tanaka M., Fox B.A., Inselburg J.;
RT "An enlarged largest subunit of Plasmodium falciparum RNA polymerase
RI II defines conserved and variable RNA polymerase domains.";
RL Nucleic Acids Res. 17:9621-9636(1989).
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
CC {RNA} (N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
CC TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST
CC COMPONENT OF RNA POLYMERASE II.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: The tandem 7 residues repeats can be highly phosphorylated.
CC The phosphorylation activates POL2.
CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases were
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
CC III for 5S and tRNA genes.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15561; CAA34560.1; -.
CC PIR; S07485; RNZQ2L.
CC InterPro; IPR007022; RNA_pol_A.
CC InterPro; IPR007080; RNA_pol_Rpb1_1.
CC InterPro; IPR007066; RNA_pol_Rpb1_3.
CC InterPro; IPR007083; RNA_pol_Rpb1_4.
CC InterPro; IPR007081; RNA_pol_Rpb1_5.
CC InterPro; IPR007075; RNA_pol_Rpb1_6.
CC InterPro; IPR007073; RNA_pol_Rpb1_7.
CC InterPro; IPR006592; RNA_pol_N.
CC InterPro; IPR00684; RNA_polII_repeat.
CC Pfam; PF04997; RNA_pol_Rpb1_1.
CC Pfam; PF00623; RNA_pol_Rpb1_2.
CC Pfam; PF04983; RNA_pol_Rpb1_3.
CC Pfam; PF05000; RNA_pol_Rpb1_4.
CC Pfam; PF04998; RNA_pol_Rpb1_5.
CC Pfam; PF04992; RNA_pol_Rpb1_6.

```

```

DR Pfam; PF04990; RNA_pol_Rpb1_7; 1.
DR Pfam; PF05001; RNA_pol_Rpb1_R; 12.
DR SMART; SM00663; RPOLA_N; 1.
DR PROSITE; PS00115; RNA_POL_II_REPEAT; 9.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN FING 68 84
FT C2H2-TYPE (POTENTIAL).
FT BY SIMILARITY.
FT DNA BIND 378 411
FT POLY-ASN.
FT DOMAIN 707 725
FT LEUCINE-ZIPPER (POTENTIAL).
FT DOMAIN 1093 1128
FT POLY-ASN.
FT DOMAIN 1144 1159
FT VERY BASIC (POTENTIAL).
FT DNA BIND 1182 1193
FT POLY-ASN.
FT DOMAIN 1258 1290
FT ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 1261 1290
FT 6 X TANDEN REPEATS OF [YLV]-D(3,4).
FT DOMAIN 1602 1612
FT ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 1745 1759
FT ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 1806 1820
FT ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 2061 2246
FT HIGHLY DIVERGED HEPTAPEPTIDE REPEATS.
FT DOMAIN 2247 2384
FT CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
SQ SEQUENCE 2452 AA; 278166 MW; F995E117F617A48F CRC64;

Query Match 6.1%; Score 183; DB 1; Length 2452;
Best Local Similarity 21.1%; Pred. No. 2.2;
Matches 145; Conservative 95; Mismatches 275; Indels 172; Gaps 28;

QY 3 DEKNDVINNNOSINTDNNQIIKKEETNNYDGIKRSSEDRTESTTNVDENEATFLQK-- 60
DB DDDLDDEYNSDININIGNRKYYGNLKNYD-----ENSMNPIDVVHKVNNFLEKLV 1336
QY 61 -TPQDNTHITEREKSSSVESNSSIDTAQPSHTTINRESVQTS-----DN 108
DB 1337 IIKQINSN-----DTLSVEAQNNATILLKAHLRTYLSKLLTOTHKVSVKGLDWLQ 1389
QY 109 VE-----DSHVSDFANSKIKESNTES-----GKEENTIQP-----NKV 142
DB 1390 IEKIFKYSCHPGECVGAALAAQSIGEPATQMTLNTFPHAGVGSKNVTGVPRLKELINIV 1449
QY 143 KEDSTTSQPSGYTNIDEKISNQDELLNLPINBYENKAREPLSTTSAPSIKRVTVNQLAAE 202
DB 1450 KNVKT--PSTTYLDDMVSNQOKAKDILTLEVTTLKQLTSHAQIIIVDPNTTITILEE 1506
QY 203 QGSNNHLIKVTDQSTEGYDSDSEGVKAKDAENLIYDVFVDDKVKSGDWTVDIDKN 262
DB 1507 DKSWNVEFYEPFDDDTQ-YSLGEWVLR-----IQLTNIHVNEK-KLTMRKEIVIIYS 1557
QY 263 TVPSDLDTFTPIKINSGEII-----ATGYDNKNKQITVTFDVIDVYENIKAKHL 315
DB 1558 VFSSDELD---LIYTDNSEDIVLRVVKYLNGEFNFMYDVVDNANEQVDSEDEEHL 1614
QY 316 -----KLTSYIDKSKVPNN-----NTK-----LDVEYKTALS 342
DB 1615 VANDRGNYDETKNSTPHPHDYNNNTTNIFKSKVKNNISSDINTKNEDSISINSSNNEQVK 1674
QY 343 SVNKITVVEYQPNENRFTANLOSMTNIDTKNHTVE-----QTIYINPLRYS 389
DB 1675 NINSPVSNMNMNNNNNNNDSSNINDIKVKNIKKEDGNEGALRGCGDSNTSALFGNKNS 1734
QY 390 AKETNV--NISGNGDEGSTIIDDSTIIKVKYKVDNQNLPSNRIVDYSEYEDVTND-- 444
DB 1735 QKEDINVNNDNNDND--DEEEEDFLGCDHNVSPKNTKDGKNKNNKNNNNENK 1790
QY 445 YAOLGNNNDVINNF--GNIDSPYIIKVISKVDPNKDDYTTTIOQTVMQTTTINEYTGEP 501
DB 1791 NKKSGNNNSNNTYDDGDVDN-----DNDDDDNDKNSDI-----TIKE----- 1829
QY 502 TASYDNTAF-STSSGQGGDLPPKTYKIGDYVWEDVDKGIQNT---NDNEKPLSNVL 557
DB 1830 -----DNDVAFMKTSTKNAEEDL-----ELKNKNHSHISREDETFLLKLMQECCLSTL 1881
QY 558 -----VTITYPDGTGSKSVRTDEGKY 578
DB 1882 LAGIENITKVYREESKITVDSNKGK 1908

```

RESULT 15
 MRC1 YEAST STANDARD; PRT; 1096 AA.
 AC P25589; P27513; P87003; Q07218; Q8NIN2;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA replication checkpoint mediator MRC1.
 GN MRC1 OR YCL061C OR YCL61C/YCL60C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=92244356; PubMed=1574125;
 RA Oliver S.G., van der Aart Q.J.M., Agostoni-Carbone M.L., Aigle M.,
 RA Alberghina L., Alexandraki D., Antoine G., Anwar R., Ballesta J.P.G.,
 RA Benit P., Berben G., Bergantino E., Biteau N., Bolle P.-A.,
 RA Bolotin-Fukuhara M., Brown A.J.P., Brown R., Buhler J.-M.,
 RA Carignani G., Chanet R., Contreras R., Crouzet M., Daigman-Fornier B.,
 RA De Haan M., Defoor E., Delgado M.D., Demolder J., Doira C., Dubois E.,
 RA Dujon B., Duesterhoeft A., Erdmann D., Estebar M., Fabre F.,
 RA Fairhead C.A., Faye G., Feldmann H., Fiers W.,
 RA Francinques-Gaillard M.-C., Franco L., Frontali L., Fukuhara H.,
 RA Fuller L.J., Gent M.E., Gigot D., Gilliquet V., Glandsdorff N.,
 RA Goffeau A., Grenson M., Grisanti P., Grivell L.A., Haaseemann M.,
 RA Hatat D., Hegemann J.H., Herbert C.J., Hilger F., Hohmann S.,
 RA Hollenberg C.P., Huse K., Iborra F., Indge K.J., Isono K., Jackman P.,
 RA Jacq C., Jaquet M., James C.M., Jauniaux J.-C., Jia Y., Jimenez A.,
 RA Kleinhaus U., Kreisel P., Lafranchi G., Lewis C., van der Linden C.G.,
 RA Lucchini G., Lutzenkirchen K., Maat C.C., Mannheim G., Manzano M.E.,
 RA Martegani E., Mathieu A., Maurer C.T.C., McConnell D., McKee R.A.,
 RA Messenguy F., Mewes H.-W., Molemans F., Montague M.A., Navas L.,
 RA Newton C.S., Olson M.V., Pallier C., Panzeri L., Pearson B.M.,
 RA Perea J., Philippsen P., Piarard A., Planta R.J., Plevani P.,
 RA Poetsch B., Pohl F.M., Purnelle B., Ramezani Rad M., Raguenes S.W.,
 RA Raynal A., Remacha M., Richterich P., Roberts A.B., Rodriguez F.,
 RA Sanz Y., Schaaff-Gerstenschlaeger I., Scherens B., Schweitzer B.,
 RA Shu Y., Skala J., Slonimski P.P., Sor F., Soustelle C.,
 RA Spiegelsberg R., Staveva L.I., Steensma H.Y., Steiner S., Thierry A.,
 RA Thiores G., Triano L.N., Urrestazu L.A., Valle G., Vetter I.,
 RA van Vliet-Reedijk J.C., Volckaert G., Vreken P., Warwington J.R.,
 RA von Wettstein D., Wickstead B.L., Wilson C., Wurst H., Xu G.,
 RA Zimmermann F.K., Sgouras J.G.
 RL "The complete DNA sequence of yeast chromosome III."
 RT Nature 357:38-46(1992).
 [2]
 RP REVISIONS.
 RA Gromadka R.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Valles G., Volckaerts G.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 660-839 FROM N.A.
 RC STRAIN=ATCC 28383 / FL100;
 RX MEDLINE=90384830; PubMed=2169608;
 RA Kern L.;
 RL "The URK1 gene of Saccharomyces cerevisiae encoding uridine kinase."
 RT Nucleic Acids Res. 18:5279-5279(1990).
 RN [5]
 RP FUNCTION, AND PHOSPHORYLATION.
 RX MEDLINE=21571912; PubMed=11715016;
 RA Alcasabas A.A., Osborn A.J., Bachant J., Hu F., Werler P.J.,
 RA Bousset K., Furuwa K., Diffley J.F., Carr A.M., Elledge S.J.;
 RT "Mrc1 transduces signals of DNA replication stress to activate
 Rad53."
 RL Nat. Cell Biol. 3:958-965(2001).

FN
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=22747592; PubMed=12865299;
 RA Osborn A.J., Elledge S.J.;
 RT "Mrc1 is a replication fork component whose phosphorylation in
 response to DNA replication stress activates Rad53."
 RL Genes Dev. 17:1755-1767(2003).
 [7]
 RP INTERACTION WITH CDC45.
 RX MEDLINE=22825747; PubMed=12944972;
 RA Katou Y., Kanoh Y., Bando M., Noguchi H., Tanaka H., Ashikari T.,
 RA Sugimoto K., Shirahige K.;
 RT "S-phase checkpoint proteins Tof1 and Mrc1 form a stable
 replication-pausing complex."
 RL Nature 424:1078-1083(2003).
 CC -!- FUNCTION: Required for normal DNA replication. Phosphorylated in
 response to DNA replication stress. Phosphorylation allows it to
 mediate the activation of RAD53.
 CC -!- SUBUNIT: Interacts with CDC45 in S phase.
 CC -!- SUBCELLULAR LOCATION: Nuclear; associated with chromatin during S
 phase.
 CC -!- PTM: Phosphorylated by MEC1 and RAD53.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X59720; CAC42953.1; -;
 DR EMBL; X53998; CAA37945.1; -;
 DR PIR; S74279; S74279.
 DR Germline; 138895; -;
 DR SGD; S0000566; MRC1.
 DR CO; GO:0006347; P:chromatin silencing at HML and HMR (sensu S. .; IGI.
 DR GO; GO:0006348; P:chromatin silencing at telomere; IGI.
 DR GO; GO:0000076; P:DNA replication checkpoint; IGI.
 DR GO; GO:0000076; P:DNA replication; Phosphorylation; Coiled coil.
 KW Nuclear protein; DNA replication; Phosphorylation; Coiled coil.
 FT DOMAIN 488 542 COILED COIL (POTENTIAL).
 FT DOMAIN 652 716 COILED COIL (POTENTIAL).
 FT DOMAIN 748 808 L -> V (IN REF. 4).
 FT CONFLICT 748 808 MISSING (IN REF. 4).
 FT CONFLICT 808 808 MISSING (IN REF. 4).
 SQ SEQUENCE 1096 AA; 124325 MW; 378345E5503FFA81 CRC64;
 Query Match 6.1%; Score 182.5; DB 1; Length 1096;
 Best Local Similarity 18.9%; Pred. No. 0.88;
 Matches 121; Conservative 117; Mismatches 234; Indels 169; Gaps 29;
 QY 5 EKNDVNNQSIINTDNNQ-----IIKKEETNNYDGIKESDRTSTNNVDNEATFLQK 60
 DB 534 EENDFQLNAHDSGSDSGSSGFGALSGNEIADYESSGSENDNRNRSDEKEDDEILKQK 593
 QY 61 TPQDNTHL-----TEEEVKSSSVSSSSNSTDYTAQPSHTTN-----REESVQSDNV 109
 DB 594 KSHVHVHIIINESDSTVEAKPKKAKDES-----LPKRIALNIGHYDNGTGEDTKF 645
 QY 110 EDHSVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSOPSGYTNIDKISNQ---DE 166
 DB 646 QETNVLDTQNI-----EVMARNTIE--NEVDDVYNEEA-----DEAIRQLIDKE 692
 QY 167 LNLPLINEYENKARPLSTTSQAQSIKRVTVN--QLAAEQGSNNVNHILIKVTDQSITEGYDD 224
 DB 693 KLQKQKEHEAK-----IKELKKRGVTNPFMEAESEDESWHGGAGGSDGSDYDS 746
 QY 225 -----SEGVIKAHDAENLIYDTFFVDKVKSGDWTVDID-----KNTVPS 266
 DB 747 DLEKMIDDYSKNFNPHREMLAENKEMD--IKMINKILYDIKNGGFRNKAKNSLEL 804
 QY 267 DLTDSFTPIPKDNGSEIATGYDNKNKQITVTFDYVDYKYNKIAHLKLTYSVSKSV 326
 DB 805 ELSD-----DDEDDVLQ--QYRLKRLMLR-----KRLKLEIG---DDAKL 839

Search completed: October 5, 2004, 20:48:56
Job time : 27.1685 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 65.362 Seconds
(without alignments)
2809.455 Million cell updates/sec

Title: US-10-806-288-13
Perfect score: 3013
Sequence: 1 SDEEKNDVNNQSIINTDD.....PDGTSKSVRTDEGKYQFDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archepa:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1056	16	Q8CQ72
2	3013	100.0	1092	2	O70022
3	2810.5	93.3	931	2	Q9K113
4	1160.5	38.5	1171	2	Q9KWX6
5	1144.5	38.0	1166	2	O86489
6	1143.5	38.0	1141	16	Q8NXX5
7	1137.5	37.8	1141	16	Q99W46
8	1137.5	37.8	1141	16	Q932F7
9	576	19.1	1315	2	O86488
10	567	18.8	1347	16	Q8NXX6
11	559	18.6	1385	16	Q99W47
12	499	16.6	1733	2	Q9K114
13	497	16.5	1633	16	Q8CWP4
14	489	16.2	953	16	Q99W48
15	485.5	16.1	1893	2	Q8KWH1
16	470.5	15.6	946	16	Q8NXX1

17	465	15.4	955	16	Q8NXX7	Q8nxx7 staphylococ
18	460.5	15.3	881	2	Q93MH7	Q93mh7 staphylococ
19	458.5	15.2	566	2	Q8KR22	Q8kr22 staphylococ
20	455	15.1	947	2	O86487	O86487 staphylococ
21	452	15.0	933	2	Q53653	Q53653 staphylococ
22	452	15.0	935	16	Q932C5	Q932c5 staphylococ
23	452	15.0	989	16	Q99VJ4	Q99vj4 staphylococ
24	397.5	13.2	877	16	Q99R07	Q99r07 staphylococ
25	396	13.1	913	2	O86476	O86476 staphylococ
26	384.5	12.8	907	16	Q8NUL0	Q8nul0 staphylococ
27	376	12.5	940	2	Q53682	Q53682 staphylococ
28	374.5	12.4	943	16	Q8NUU8	Q8nuu8 staphylococ
29	344.5	11.4	961	16	Q99RD3	Q99rd3 staphylococ
30	326.5	10.8	1015	16	Q8NUU7	Q8nuu7 staphylococ
31	326	10.8	1038	16	Q99RD2	Q99rd2 staphylococ
32	293.5	9.7	970	16	Q8DYL7	Q8dy17 streptococ
33	293.5	9.7	1310	16	Q8E473	Q8e473 streptococ
34	276	9.2	2616	5	Q8IIG1	Q8iig1 plasmodium
35	268.5	8.9	1301	5	Q8WSK5	Q8wsk5 plasmodium
36	266.5	8.8	1160	2	Q8RM86	Q8rm86 streptococ
37	266.5	8.8	1301	5	Q8IHQ2	Q8ihq2 plasmodium
38	266	8.8	2402	2	Q9AER7	Q9aer7 staphylococ
39	265	8.8	3452	5	Q8IEA3	Q8iea3 plasmodium
40	261	8.7	1161	2	Q9X3M7	Q9x3m7 streptococ
41	261	8.7	2153	5	Q8ICM9	Q8icm9 plasmodium
42	259	8.6	1161	2	Q8RJL0	Q8rjl0 streptococ
43	253.5	8.4	3078	5	Q8IKH4	Q8ikh4 plasmodium
44	253.5	8.4	3207	5	Q8IHQ0	Q8ihq0 plasmodium
45	249.5	8.3	1859	5	Q8IC27	Q8ic27 plasmodium

ALIGNMENTS

RESULT 1

Q8CQ72	PRELIMINARY;	PRT; 1056 AA.
ID	Q8CQ72	
AC	Q8CQ72;	
DT	01-MAR-2003 (TREMBLrel. 23, Created)	
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)	
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DE	Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein.	
DE	SE0331.	
GN	Staphylococcus epidermidis.	
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.	
OX	NCBI_TaxID=1282;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 12228;	
RA	Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z., Chen Z., Wen Y.;	
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; A5016745; AAC03928.1; -	
DR	GO; GO:0015020; Cmembrane; IEA.	
DR	InterPro; IPR005877; Gpos_Ysirk.	
DR	InterPro; IPR001899; Gram_pos_anchor.	
DR	Pfam; PF05738; Cna_B; 2.	
DR	Pfam; PF00746; Gram_pos_anchor; 1.	
DR	Pfam; PF04650; Ysirk_signal; 1.	
DR	TIGRFAMs; TIGR01167; LPXTG_anchor; 1.	
DR	TIGRFAMs; TIGR01168; Ysirk_signal; 1.	
DR	PROSITE; P550847; GRAM_POS_ANCHORING; 1.	
KW	Complete proteome.	
SQ	SEQUENCE 1056 AA; 115728 MW; 015869A9B5CA2723 CRC64;	

Query Match 100.0%; Score 3013; DB 16; Length 1056;
Best Local Similarity 100.0%; Pred. No. 2.5e-108;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDEEKNDVNNQSIINTDDNQIIKKETNNYDGIKESDEPTSTNNVDENATFLOK 60
|||||


```
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; GpoS_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 2; Gaps 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 93.3%; Score 2810.5; DB 2; Length 931;
Best Local Similarity 93.6%; Pred. No. 1.3e-100;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEKNDVNNQSIINTDDNQIIKKEETNNYDGIEKRSDETESTTNDNEATFLOK 60
DB 74 SSNEKNDVNNQSIINTDDNQ-IKKEETNSDAIENRSKDTQSTTNDNEATFLOK 132
QY 61 TPQDNTLHTEEEVKESSVSSSSIDTAQPSHTTINREESVQTSNDVEDSHVDFANS 120
DB 133 TPQDNTQLKEEVKESSVSSSSMDTAQPSHTTINSEASQTSDNEENSRVDFANS 192
QY 121 KIKESNTESGKEENTIEQPNKVEDSTSPSGYNTIDEXISNQDELLNLPINEYENKAR 180
DB 193 KIKESNTESKEENTIEQPNKVEDSITSQSSYKNIDEXISNQDELLNLPINEYENKVR 252
QY 181 PLSTTSQPSIKRVTVNQLAABQGSNNHLIKVTDQSITEGYDDSGEVIKAHDAENLIYD 240
DB 253 PLSTTSQPSKRVTVNQLAABQGSNNHLIKVTDQSITEGYDDSGEVIKAHDAENLIYD 312
QY 241 VTEVEDDKVKSQDGTMTVDIDKNTVPDLTDSFTPIKIKNSGIIATGTYDKNKQIITYT 300
DB 313 VTEVEDDKVKSQDGTMTVDIDKNTVPDLTDSFALPKIKNSGIIATGTYDNTNKQIITYT 372
QY 301 FTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALSSVKNKTIITVEYQRPNEKRT 360
DB 373 FTDYVDKYENIKAKHLKLTYSIDSKVPNNNTKLDVEYKTALSSVKNKTIITVEYQKPNERT 432
QY 361 ANLQSMFTNIDTKNHTVEQTIYINPLAYSARETNVINSNGDESGTIIIDSTIIKYKVG 420
DB 433 ANLQSMFTNIDTKNHTVEQTIYINPLAYSARETNVINSNGDESGTIIIDSTIIKYKVG 492
QY 421 DNQNLPSNRIDYSEYEDVTNDYAOLGNNNDVNFNGNIDSPYIIKVLISKYDPNKDDY 480
DB 493 DNQNLPSNRIDYSEYEDVTNDYAOLGNNNDVNFNGNIDSPYIIKVLISKYDPNKDDY 552
QY 481 TTIQQTVMQTINEYTGERTASYDNTIAFSTSSGQGGDLPPKTYKIGDYVWEDVDK 540
DB 553 TTIQQTVMQTINEYTGERTASYDNTIAFSTSSGQGGDLPPKTYKIGDYVWEDVDK 612
QY 541 DGIONTNDNEKPLSNVLVTIYDPDGTSKSVRTDEDKYQPDG 582
DB 613 DGIONTNDNEKPLSNVLVTIYDPDGTSKSVRTDEEGKYQPDG 654

RESULT 4
Q9KWX6 PRELIMINARY; PRT; 1171 AA.
AC Q9KWX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bone sialoprotein-binding protein.
GN BBP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.

STRAIN=024;
MEDLINE=20115096; PubMed=10642520;
Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
"A bone sialoprotein-binding protein from Staphylococcus aureus: a
member of the staphylococcal Sdr family 1.";
Biochem. J. 345:611-619(2000).
RL Biochem. J. 345:611-619(2000).
CC 1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
DR EMBL; Y18653; CAB75732.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; GpoS_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1171 AA; 127123 MW; CSBC812F9DA5A884 CRC64;

Query Match 38.5%; Score 1160.5; DB 2; Length 1171;
Best Local Similarity 42.9%; Pred. No. 4e-37;
Matches 271; Conservative 99; Mismatches 182; Indels 79; Gaps 20;

QY 1 SSDEKND---VINNQSIINTDDNQIIKKEETN-----NYDGIKRSDETESTTND 51
DB 57 STENAKQDEASADNKEVVSETENNSTQKNDLTNPICKETNTDHSQEAAPTSTSTQQQ 116
QY 52 ENEATFLQTPQDNTLHTEEEVKESSVSSSSIDTAQPSHTTINREESVQTSND 108
DB 117 QNNATTSTETQPN--IEKENVKPSTDKATDTSVILEEKKAPNNT--NNDVTTKPS-- 170
QY 109 VEDSHVSDPANSKIKESNT---ESGKEENTIEQPNKVED---STTSQPSGYNTI--DEK 160
DB 171 -----TSIQITPTTPQESTNIENSQOPTSKVDNQVDTDTNPKPEPVNSKEEL 220
QY 161 ISNQDELLNLPINE--YENKARPLSTTSQPSIKRV-----TVNQLAABQGSNNVHLIK 212
DB 221 KKNPEKLKELVRNDSNTDRSTKPVATPTSPVAKRVNAKIRFAVAQPAVASNNVDLIT 280
QY 213 VTDQSITEGYDDSGEVIKAHDAENLIYDVTFFEDVKVSGDNTVTDIDKNTVPSDLTDSF 272
DB 281 VTKQMITTEGIKD-DGVIQAHDEGHIITYSDFKIDNAVAGADTMTVKYKETIPESDITDDF 339
QY 273 TIPKIKDNGSEIITATCTYDKNKQIITYTFTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTK 332
DB 340 TPVDITDPSGEVJAGTFFDLNTKIITYKFTDYVDKYENNAKLELSYIDKKEVP-NETN 398
QY 333 LDVEYKTALSSVKNKTIITVEYQRPNEKRTANLQSMFTNIDTKNHTVEQTIYINPLAYSARE 392
DB 399 LNLTFATADKETSNNKVEYQKPIVKDESNIQISFSLDITTKHEVEQTIYVNPVNLKNAKN 458
QY 393 TNVNI-----SGNGD-----EGSTIIDSTIIKYKVGDNQNLPSNRIDYSEYEDVTND 443
DB 459 TNVTIKSGGVADNGDYITDGTSTIIDSTNTEIKYKVASGQQLPSQNKIYDYSQYEDVTNS 518
QY 444 DYLAQLGNN--NDVNINFGNIDSPYIIKVLISKYDPNKDDYTTTIQQTVMQTINEYTGFE 500
DB 519 --VTINKNYGTNMANINFGDIDSAIYKVVSKITPGAEDDLAVQQGVMTT-----TNKY 571
QY 501 RTASYDN-----TIAFSTSSGQGGDLPPKTYKIGDYVWEDVDKDGIONTNDNEK 551
DB 572 NYSSYAGYITLFYQLLTLVVVTVSVK-----PEEKLYKIGDYVWEDVDKGVQGTDSKEK 626
QY 552 PLSNVLVTIYDPDGTSKSVRTDEDKYQPDG 582
DB 627 PMANVLVTIYDPDGTTSKSVRTDANGHYEFGG 657

RESULT 5
```

```
086489
ID O86489 PRELIMINARY; PRT; 1166 AA.
AC
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Sdr E protein.
GN SDR E.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Newman;
RC MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
DR EMBL; AJ005647; CAA06652.1; -
DR PIR; T28680.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; peptidoglycan-anchor.
SQ SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;

Query Match 38.0%; Score 1144.5; DB 2; Length 1166;
Best Local Similarity 42.6%; Pred. No. 1.6e-36;
Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;

QY 1 SSDEKNDVINNQSIINTDNNQI--IKKETNNYDGIKESDRSTSTNVNDEATFL 58
DB 66 ATTSNKEVSEVENSTENNSTNP1KGE--TNTDSQPEAKEESTSSSTQKQNNVTAT 123
QY 59 QKTPODTHLTEREVKES---SSVESNSSIDTAQPSHTTINKEESVQTSNDVDSHVS 115
DB 124 TETKPN--IEKENVKSTDKTATEDSVILEEKKAPNT--NNDVTTPS-----TS 172
QY 116 DFANSKIKESNT---ESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKIS--NODELLNLP 171
DB 173 EPSTSEIQTKPTTPESTNIENSQPQTPSKVD---NQVTDATNPKEPVNVSKELKNP 229
QY 172 -----INEYENKARPLSTSAQSIRKV-----TVNQLAAGSGSNVNHILKVTQ 216
DB 230 EKULVRNSDTHSTKPVATAPTSVAPKRVNAKMFVAQAAPAAVANNVNDLIKVTQ 289
QY 217 SITEGYDDSGVKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPK 276
DB 290 TIKVG-DGKNVAHAHDGKDIEYDTEFTIDNKVKGDTMTINTYDKNVPSDLTKNDPID 348
QY 277 IKNSGBIIATGYDNNKQKQITVFTFDYDVKYENIKAHKLKLTYSIDKSKVPNNNTKLDVE 336
DB 349 ITDPSGEVIAKGTFDKATKQITVFTFDYDVKYEDIKSLRTLSYIDKKTVP-NETSLNLT 407
QY 337 YKTALSSVNKTIIVEYORPENENTANLQSMFTNIDTKHVEQTIYINPLRYSAKETNVN 396
DB 408 FATAGKETSQNTVVDQPMVHGDSNIQSIFTKLBDEKQIEQCIYVNPPLKKSATNTKVD 467
QY 397 ISGNG-----DEGSTIIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTND-DYA 446
DB 468 IAGSQVDYGNIKLNGSTIIDQNTTEIKVKYKVSQDLQPSNRIYDFSYEDVTSQFDNK 527

447 QLNQNDVNFNGNIDSPYIIKISKYDPNKKDYTTIQOVTVMQTTINEYTGFRRTASYD 506
528 KSFNNVATLDFGDIINSAYIIKVKYTPFSDGELDIAQCTSMRTT-DKY-GYNTAGYS 585
507 NTIAFTSTSGQGGLD-LPEKTYKIGDYVWEDVDKGIQNTNDNERKPLSNVLTLYTPDG 565
586 NFIVTSNDTGGDGTVPKEKLYKIGDYVWEDVDKGVQGTDSKEKPMANVLTLYTPDG 645
566 TSKSVRTDEGKYQFDG 582
646 TTKSVRTDANGHYEFG 662

RESULT 6
Q8NXX5 PRELIMINARY; PRT; 1141 AA.
ID Q8NXX5
AC Q8NXX5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Sar-Asp rich fibrinogen-binding bone sialoprotein-binding
DE protein.
GN SDR OR MW0518.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AF004823; BAB94383.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1141 AA; 123997 MW; 372E5860850A332C CRC64;

Query Match 38.0%; Score 1143.5; DB 16; Length 1141;
Best Local Similarity 43.3%; Pred. No. 1.8e-36;
Matches 266; Conservative 98; Mismatches 192; Indels 59; Gaps 18;

QY 1 SSDEKNDVINNQSIINTDNNQI--IKKETNNYDGIKESDRSTSTNVNDEATFL 60
DB 69 SDNKEVSEVENSTENNSTNP1--KKTNT--DSQPEAKEESTKSSTQKQNNVTATTE 125
QY 61 TPQDTHLTEREVKES---SSVESNSSIDTAQPSHTTINKEESVQTSNDVDSHVSDF 117
DB 126 TTPQN--IEKENVKSTDKTATEDSVILEEKKAPNT--NNDVTTPS----- 170
QY 118 ANSKI KESNT---ESGKEENTIEQPNKVEDSTTSQPSGYTNIDEKIS--NODELLNLP-- 171
DB 171 -TSEIQTKPTTPESTNIENSQPQTPSKVD---NQVTDATNPKEPVNVSKELKNP 226
QY 172 -----INEYENKARPLSTSAQSIRKV-----TVNQLAAGSGSNVNHILKVTQSI 218
DB 227 LKELVRNSDTHSTKPVATAPTSVAPKRVNAKMFVAQAAPAAVANNVNDLIKVTQTI 286
QY 219 TEGYDDSGVKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPK 278
DB 287 KVG-DGKNVAHAHDGKDIEYDTEFTIDNKVKGDTMTINTYDKNVPSDLTKNDPIDIT 345
```



```

DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF05738; Cna_B; 3.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YSIRK signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK signal; 1.
DR PROSITE: PS0847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1141 AA; 124038 MW; E679F7C2991846D9 CRC64;

Query Match
Best Local Similarity 37.8%; Score 1137.5; DB 16; Length 1141;
Matches 266; Conservative 97; Mismatches 193; Indels 59; Gaps 19;

QY 1 SSDEKNDVNNOSINTDNNQIIKKEETNNYDGIKRSDETRTESTTNVDENAEFLQK 60
DB 69 SDNKEVSVETNNSTENDSNPI--KKEINT--DSQPEAKEESTTSQQQNNVTATTE 125
QY 61 TPQDNTHLTREEVKES---SSVESNSSIDTAQOPSHHTINREESVQTSNVEDSHVSD 117
DB 126 TKPQN--IEKENVPSTDKTATEDTSVLEKAPNYT--NNDVTTKPS----- 170
QY 118 ANSRIKESNT---BSGKEENTIEQPNKVEDSTTSQPSGYTNIDKIS--NODELLNLP-- 171
DB 171 -TSEIQKPTTPQESTNIENSQOPTESKD---NQVTDATNPKPEVNVSKKEELNNPEK 226
QY 172 -----INEYENKARPLST--TSAQP-----SIKRVTVNQLAARQGSNNVNHILKVTDSQI 218
DB 227 LKELVRDNNNTDRSTKPVATAPTSVAPKRLNAKMPFAVAPAAVASNNVDLITVTKQTI 286
QY 219 TEGYDDSEGVKHADEANLIYDVTFFVDDKVKSGDWTVDIDKNTVPSDLTDSFTPIPKIK 278
DB 287 KVG-DGKDNVAAAHGDKDIEYDTEFTIDNKVKGGDTWTINYNKKNVIPSDLTKNDPIDIT 345
QY 279 DNSGEIATGYDNKNKQITVFTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYK 338
DB 346 DPSEVIKAGTFDKATKQIYFTDYVDKYEDIKARLTLSYIDKQAVP-NEISLNLTFA 404
QY 339 TALSSVNTITVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNIS 398
DB 405 TAGKETSONSVQYDPMVHGDSNIQSIFTKLDENKQTIIEQIYVNPFLKKTATNTKVDA 464
QY 399 GNG-----DEGSTIIDSTIIKVKYKQDNQNLPSNRIDYSEYEDVTND-DYAQL 448
DB 465 GSQVDDYGNIKLNGSGTIIDONTEIKVKYKPNQQLPQSNRIYDFSQYEDVTSQFNKKS 524
QY 449 GNNNDVNFNGNIDSPYIIKVISKYPDNKDYYTTIQOTVTMTTINBYTGEFTASVDNT 508
DB 525 FSNNAVLDFGDNINSALIIKVSQKPTSDGELDIAGTSMRTT-DKY-GYNNVAGYSNF 582
QY 509 IAFSTSSGGQGGDL-PPEKTYKIGDYVWEDVKDGIQNTNDNKPPLSNVLVTLTYPDGTS 567
DB 583 IVTSNDTGGGDTGVKPEKLYKIGDYVWEDVKDGVCGQTSDEKPMANVLVTLTYPDGTT 642
QY 568 KSVRTDEDEGKYQPDG 582
DB 643 KSVRTDANGHYEFGG 657

RESULT 9
ID O86488 PRELIMINARY; PRT; 1315 AA.
AC O86488;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Srd protein.
GN SRD.
OC Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=Newman;
RA MEDLINE=99098700; PubMed=9884231;
RA Joseason E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
CC -|- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL: AJ005646; CAA06651.1; -.
DR PIR: T28679; T28679.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR005877; Gpos_YSIRK.
DR Pfam: PF05738; Cna_B; 5.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YSIRK signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK signal; 1.
DR PROSITE: PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1315 AA; 142775 MW; 58D4E1F48EE6A689 CRC64;

Query Match
Best Local Similarity 30.1%; Pred. No. 1.3e-14;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 1 SSDEKNDV---INNQSINTDNNQIIKKEETNNYDGIKRSDETRTESTTNVDE-NEAT 56
DB 55 STNKELNEATTSASDNQSSDKVDMQQLNQEDNTKNDNQEMVSSQGNETTSNGKLIKE 114
QY 57 FLQKTPQDNTHLTREEVKESSSVESNSSIDTAQOPSHHTINREESVQTSNVEDSHVSD 116
DB 115 SVQSTTGKVKFVSTAKSDEQASPKSTNEDLNTKQ-----TISNQELQ-PDIQENKSVN 168
QY 117 FANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNI-DEKISNODELL--NLPI 173
DB 169 -----VQPTNEENKQVD-----AKTESTT-----LNVKSDAIKSNDETLVDNNS 209
QY 174 EYENKARPLSTSAQP-----SIKRVTVNQLAARQGSNNVNHILKVTDSQITEGYDDSEGV 229
DB 210 NNENNADIILPKSTAPKRLNTRMRIAQVPSSTEAQKNVNDLITSNTLTVDADKNNKIV 269
QY 230 KAHDAENLIYDVTFFVDDKVKSGDWTVDIDKNTV-----PSDLTDSFTPIKID-NSG 282
DB 270 PAQDYLKSKQIT--VDDKVKSGDYFTIKY-SDTVQVYGLNPFEDIKN---IGDIKDPNNG 323
QY 283 EIIATGYDNKNKQIYFTDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALS 342
DB 324 ETIATAKHDTANNLITYFTDYVDRFNSVQMGINSIYMDADTIP--VSKNDVEFNVTIG 381
QY 343 SVNKITITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYINPLRYSAKE 392
DB 382 NTTTKTTANIQYDPYVNEKNSIGSAFT--ETVSHVGNKENFGYKQTIYVNPENSELTN 439
QY 393 TNVNI-----SGNGDEGSTIIDSTIIKVKYKQDNQNLPSNRIDYVS--EYEDVTNDVYA 446
DB 440 AKLKQVQYHSSYPNNGIQINKQVTDIKIYQVPGYTL---NKGVDNTEKELTDVTN-QYL 495
QY 447 Q---LGNNDVNFNGNIDSPYIIKVISKYPDNKDYYTTIQOTVTMTTINBYTGEFTFA 503
DB 496 QKITTYGDNNSAVIDFGNADSAVVMVNTKFTQNTSESPTLVQWATLSSTGN-----KSV 549
QY 504 SYDNTIAFSTSSGGQGGDL-PPEKTYKIGDYVWEDVKDGIQNTNDNKPPLSNVLVTLTY 563
DB 550 STGNALGFTNNQSGAG-----QEVYKIGNYVWEDTKNGVQEL--GEKGVNVTVTV-FD 602
QY 564 DGTSKSVR---TDEDGKY 578
DB 603 NNTNTKVGAEVTKEDGSY 620

```



```

QY 176 ENKARPLSTTGAQPSIKKVTNQLAEOGS-----NVNHLIKVTDQSTIGVDDSEGVIKA 231
DB 212 ENNADIILPKSTAPKSLNTRMAAIOPNSTDSKNVNDLITNTTLTVVDADNSKTIQVPA 271
QY 232 HDAENLIYDVTFEVDKVKSGDWTVDIDKNTV-----PSDLTDSFTIPKID-NSGEI 284
DB 272 QDYLSLKSOIT--VDDKVKSGDYFTIKY-SDTQVYGLNPEDIKN---IGDIKDPNNGET 325
QY 285 IATGYDNMKNKOITFTFYDYDYKYENIKAHKLKLSYIDKSKVPNNNTKLDVVEYKLTALSSV 344
DB 326 IATAKHDTANNLIITYFTDYDYDRENSVKMGINSYIYMDADTIPVD--KQDVPFVSFTIGNQ 383
QY 345 NKTIIVYQRP--NENRTANLQSMFTNIDTKNHT-----VQTYIINPLRYSAKETN 394
DB 384 ITTTTADITYPAYKEADNNSGSAFT--ETVSHVGNVEDPGYVQVYVVPNMDKDLKGAK 441
QY 395 VNISNGDEGST---IIDDSTIIKVKYKVGDNQNLPSNRLYDY--SEYEDVND--DYA 446
DB 442 LKVEAYHPKYPTNIGQINQNVNIIKIRVPEGYTL---NKGVDVNTDLVDVDFEKNKM 498
QY 447 QLGNNNDVNINFGNIDSPYIIKIVISKYDPNKKDDYTTIQQVTMTQTTINEYTGFEFTASYD 506
DB 499 TYGSNQSVNLDFDITSAYVVMVNTKQYTNSESPTLVQMATLSSTGN-----KSVSTG 552
QY 507 NTIAFSTSSGQGGDLPEPKYKIGDYVWEDVDKQGLQNTNDNEKPLSNVLVLTYPDGT 566
DB 553 NALGFTNNQSGAG---QEVYKIGNYWEDTNKNGVQEL--GEKGVGNVTIV-FDNNT 605
QY 567 SKSVR---TDEDGKY 578
DB 606 NTKVGEAVTKEDGSY 620

```

RESULT 12

```

Q9K114 ID Q9K114 PRELIMINARY; PRT; 1733 AA.
AC Q9K114;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Putative cell-surface adhesin SdrF.
GN SDRF.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9491;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RL Microbiology 146:1535-1546 (2000).
DR EMBL: AF245041; AAF72509.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR005877; Gpos_YsIRK.
DR Pfam: PF05738; Cna_B; 4.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YsIRK_signal; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
SQ SEQUENCE 1733 AA; 184720 MW; DBD62EA1692FD4E8 CRC64;

```

```

Query Match 16.6%; Score 499; DB 2; Length 1733;
Best Local Similarity 27.1%; Pred. No. 1.5e-11;
Matches 168; Conservative 104; Mismatches 275; Indels 72; Gaps 25;
QY 1 SDEEKDVLNNQSNINTDNNQIIKKKEFNNDYDIEKSEDEFTSTNVDENEA-TFLQ 59
DB 148 TSTTQDDSTKKNPNSLNDNLNSSLSTTSKESKTDHSTKQAMQST-NKSNLDITNDSPTQSE 206

```

```

QY 60 KT-PQDNTHLTEREVKSSSVSSNS-----SIDTAQQPSHTTIN-REESVOTSDN 108
DB 207 KTSQANNDSTDNQAPSQKLDQSKPSEQKYKTKFENDEPTQDVEHTTTLKLTFSVSTDSS 266
QY 109 VESHSVSDFPANSKIKESNTESGKEENIIQPNKVKESDSTTSQPSGYNTIDEKI----- 161
DB 267 VNDK--QDYTRSAY-----ASLGVDSNETEAITNAVRDNLDLKAASREQINEAIAEALKK 320
QY 162 --SNQDELLNLP--INEYENKARPLSTTSAQPSIKKVTNQLAEE--QGSNVNHLIKVTD 215
DB 321 DFNPDYGVDTPLALNRSQSKSP--HKSASP--RNLMSLAEPNSGKNVNDKVKITN 375
QY 216 Q--SITEYDDSGVKAHDAENLIYDVTFEVDKVKSGDWTVDIDKNTVPSDLTDSFT 273
DB 376 PTLNLKSNHANNVWPTSNQFNLKANYELDDSIKEGDTFTIKYQYIRPGGLELPAI 435
QY 274 IPKIKNSGHEIATGYTDNKNKQITFTFYDYDYKYENIKAHKLKLSYIDKSKVPNNNTKL 333
DB 436 KTLRSKQDSIVANGVYDKTNTTNTTFTYTYDYQYQNTITGSPDLIATPKRETAIKDNQNY 495
QY 334 DVEYKLTALSSVNTKIIVYQRPNNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKET 393
DB 496 PHEVTIANEVVKDFIVDYGNNKDNITT---AAVANVDNVNKHNEVVLNQNQPKYA 552
QY 394 NVNISNGDEGSTIIDSTI---IKYKVGDNQNLPS-NRIYDYSEYEDVNDYDQAL- 448
DB 553 KY-----FSTVKNGEPIGVEKYVEVTDNAMDVSFNPDLNSSLNVKDVTSQFAPKVS 604
QY 449 GNNNDVNINFGNI---DSPYIIKIVISKYDPNKKDDYTTIQQVTMTQTTINEYTGFEFTASY 505
DB 605 ADGTRVDIINFARSMANGKYYITQAVRPTGTGNVYT--EYWLTRDGTIN--TNDFYRGTK 660
QY 506 DNTIAFSTSSGQGGDLPEPKYKIGDYVWEDVDKQGLQNTNDNEKPLSNVLVLTYPDG 565
DB 661 SITVTYILNGSSTAQGNP---TVSLGDIYVLDKKNKGVO--DDDEKGLAGVYVTL--KDS 713
QY 566 TSKS-----VTDDEGKYQFD 581
DB 714 NNRELQRTVTTDQSGHYQFD 732

```

RESULT 13

```

Q8CMP4 ID Q8CMP4 PRELIMINARY; PRT; 1633 AA.
AC Q8CMP4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
GN SE2395.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AS016752; AAO06038.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR005877; Gpos_YsIRK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF05738; Cna_B; 4.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YsIRK_signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsIRK_signal; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1633 AA; 174768 MW; DAA8537B8CD289BD CRC64;

```

```
Query Match      16.5%; Score 497; DB 16; Length 1633;
Best Local Similarity 27.0%; Pred. No. 1.7e-11;
Matches 167; Conservative 105; Mismatches 275; Indels 72; Gaps 25;

QY 1 SSDEKNDVNNOSINTDDNNQIKKEETNNYDGGIEKSEDRTESTTNDENEATFQ 59
DB 148 TSTQOOSTEKNNPSLKDNLNSSTTSKESKTDHSTKQAMST-NKSLDNDSTQSE 206

QY 60 KT-PQDNTLHTEBEVKSESSVESNS-----SIDTAQPSHTTIN-REESVOTSDN 108
DB 207 KTSQANDSTNQSAKQSDSKPSEQKYKTKFENDEPTQDVEHTTKLTKTSISTDSS 266

QY 109 VEDSHVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQSPGYNTIDEKI----- 161
DB 267 VNDK--QDYTRSAV----ASLGVSDSETEAITNAVRDNLDLKAASREQINEAIIAEALKK 320

QY 162 --SNQDELLNLP--INEYENKARPLSTSAQPSIKRVTVNQLAAE--QGSNNHLKVID 215
DB 321 DFENPDYGVDTPLALNTSQSKNP--HKSASP---RNLMSLAAEPNSGKNVNDKVKITN 375

QY 216 Q--SITEGVDSSEGVKAHDAENLIYDVTPEVDKVKSGDTMTVDIDKNTVPSDLTDSFT 273
DB 376 PTLNLKSNHANNVIWPTSNQFNLKANYELDDSIKEDGTFIKYQYIRPGLELPAI 435

QY 274 IPKIDNSGELIATGYDNNKQKQITFTFYDVKYENIKAKHLKLTSYIDSKVPNNNTKL 333
DB 436 KTLQSKGDSIVANGVYDKTNTTFTFYDVKYQNTIGSFOLIAATPKRETAIKDNQNY 495

QY 334 DVEYKTLASLVNKTITVEYQRPNEENTANLOSFTIDTKHTVEQTIYINPLRYSAKET 393
DB 496 PMEVTIANEVKKDFIVDYGKNDKNTT---AAVANVDNNKHNEVYLNNQNPKYA 552

QY 394 NVNISGNGEGSTIIDSTI---IKVKVGDNQNLPS-NRIYDSEYEDVTNDYLAQL- 448
DB 553 KY-----FSTVKNGKFIPEGVKYVEYDITWAVDSENPDLNSSNVKDVTSQFTPKYS 604

QY 449 GNNNDVNIINGNI---DSPYIIKISKYDNPDKDDYTIQOQVTWQTTINEYGEFTASY 505
DB 605 ADGTRVDIINPARSWANGKYIVQAVRPTGNGVYT--EYWLTRDGTN--TNDFYRGTK 660

QY 506 DNTIASTSGQGQGLPPEKTYKIGDYWEDVDKGIQNTNDNEKPLSNVLTLYTPDG 565
DB 661 STTIVTLNGSSTAQGNP---TSLGDYVWLDKNGKVQ--DDDEKGLAGVYVTL--KDS 713

QY 566 TSKS---VRTDEDKYQFD 581
DB 714 NNRELQRTVTDQSGHYQFD 732

RESULT 14
Q99W48 ID Q99W48 PRELIMINARY; PRT; 953 AA.
AC Q99W48 ID Q99W48 PRELIMINARY; PRT; 1893 AA.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
DE protein.
GN SDRP OR SAV0561 OR SA0519.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:33:49 ; Search time 122.638 Seconds
(without alignments)
2515.875 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKQNLTKKKPIANKSN.....FAGLGALLGKKRKKNKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003Bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 AAW41602	Aaw41602 Staphyloc
2	5646	100.0	1092	7 ABM79019	Abm79019 Staphyloc
3	4490	79.5	930	5 ABP40469	Abp40469 Staphyloc
4	4485	79.4	991	3 AAY83171	Aay83171 Cell wall
5	4485	79.4	991	3 AAY70120	Aay70120 Staph. ep
6	4307	76.3	892	6 ABU42557	Abu42557 Protein e
7	3351.5	59.4	670	6 ABU42520	Abu42520 Protein e
8	2787	49.4	549	7 ABM79020	Abm79020 Staphyloc
9	2698.5	47.8	1166	2 AAY08643	Aay08643 S. aureus
10	2698.5	47.8	1166	6 ABJ18982	Abj18982 Pathogen
11	2624.5	46.5	560	7 ABM79015	Abm79015 Staphyloc
12	2586	45.8	1141	6 ABU42327	Abu42327 Protein e
13	2339.5	41.4	1633	6 ABU42513	Abu42513 Protein e
14	2303.5	40.8	1802	3 AAY83170	Aay83170 Cell wall
15	2303.5	40.8	1802	3 AAY70119	Aay70119 Staph. ep
16	2256.5	40.0	995	6 ABM72437	Abm72437 Staphyloc
17	2234	39.6	1385	6 ABU16400	Abu16400 Protein e
18	2205.5	39.1	1920	6 ABU43489	Abu43489 Protein e
19	2093.5	37.1	953	6 ABU16533	Abu16533 Protein e
20	2070	36.7	1349	6 ABM72436	Abm72436 Staphyloc
21	2069	36.6	1349	4 AAU37544	Aau37544 Staphyloc
22	2069	36.6	1349	4 AAU34402	Aau34402 Staphyloc
23	2032.5	36.0	947	6 ABJ18940	Abj18940 Pathogen
24	2011	35.6	932	4 AAU36845	Aau36845 Staphyloc
25	2011	35.6	932	4 AAU34082	Aau34082 Staphyloc

ALIGNMENTS

RESULT 1

AAW41602	1	AAW41602 standard; protein; 1092 AA.	
XX	XX	AAW41602;	
XX	XX	17-OCT-2003 (revised)	
DT	DT	22-JUN-1998 (first entry)	
XX	XX	Staphylococcus epidermidis fibrinogen binding protein FIG.	
DE	DE	Fibrinogen binding protein; FIG; aggregation; infection;	
KW	KW	coagulase-negative Staphylococcus; therapy; diagnosis; immunisation;	
KW	KW	immunogen; vaccine.	
XX	XX	Staphylococcus epidermidis; strain HB.	
OS	OS	Location/Qualifiers	
XX	XX	Key	
PH	PH	Peptide	1..51
FT	FT	/label= Sig_peptide	52..1092
FT	FT	Protein	/label= Mat_protein
FT	FT	Region	52..824
FT	FT	/note= "non-repetitive region, harbours fibrinogen binding activity"	825..1040
FT	FT	Region	/note= "Asp-Ser dipeptide repeat region"
FT	FT	Region	1053..1057
FT	FT	/note= "cell wall anchoring motif"	
XX	XX	WO9748727-A1.	
PN	PN	24-DEC-1997.	
XX	XX	18-JUN-1997; 97WO-SE001091.	
XX	XX	20-JUN-1996; 96SE-00002496.	
PA	PA	(GUSS/) GUSS B.	
PA	PA	(NILS/) NILSSON M.	
PA	PA	(FRYK/) FRYKBERG L.	
PA	PA	(FLOC/) FLOCK J.	
PA	PA	(LIND/) LINDBERG M.	
XX	XX	Guus B, Nilsson M, Frykberg L, Flock J, Lindberg M;	
XX	XX	WPI; 1998-063079/06.	
DR	DR	N-PSDB; AAV04279.	

XX Fibrinogen-binding protein from coagulase-negative Staphylococcus - used
PT for prevention, treatment and diagnosis of Staphylococcus infection.

XX Example 3; Fig 6; 45pp; English.

XX The protein comprises the fibrinogen binding protein (FIG) of coagulase-
CC negative Staphylococcus epidermidis HB. Its amino acid sequence was
CC deduced from the isolated fig gene (see AAV04279). The closest known
CC analogue of FIG is the clumping factor of *S. aureus* which also binds
CC fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG
CC polypeptides can be expressed in host cells. They are used as immunogens,
CC particularly in vaccines (which may be expressed in vivo) to protect
CC humans and animals against coagulase-negative Staphylococcus infection.
CC Antibodies raised against FIG can be used for passive immunisation. They
CC block the adherence of bacteria) and for diagnosis. (Updated on 17-OCT-
CC 2003 to standardise OS field)

XX Sequence 1092 AA;

Query Match 100.0%; Score 5646; DB 2; Length 1092;
Best Local Similarity 100.0%; Pred. No. 1.3e-266;
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAKABENSVDVK 60
DB 1 MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAKABENSVDVK 60
QY 61 DSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESDRTES 120
DB 61 DSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESDRTES 120
QY 121 TTNVDENEATFLOKTPQDNTHLTFEEVKESSESSSSIDTAQPSHTTINREESVQTS 180
DB 121 TTNVDENEATFLOKTPQDNTHLTFEEVKESSESSSSIDTAQPSHTTINREESVQTS 180
QY 181 DNVEDSHVSPANSKIKESNTESGKEENTIEQPNKVKEDSTTQPSGYTNIDKISNQDE 240
DB 181 DNVEDSHVSPANSKIKESNTESGKEENTIEQPNKVKEDSTTQPSGYTNIDKISNQDE 240
QY 241 LLNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSIITEGVDDSE 300
DB 241 LLNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSIITEGVDDSE 300
QY 301 GVIAKHAENLIYDVTFEVDVKVSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIIA 360
DB 301 GVIAKHAENLIYDVTFEVDVKVSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIIA 360
QY 361 TGYDKNKQITFTDYVDKYENIKAKHLKTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
DB 361 TGYDKNKQITFTDYVDKYENIKAKHLKTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
QY 421 TITVEYORPNERTANLQSFMTNIDTKHNTVEOTIYINPLRYSAKETNVIISNGDEGST 480
DB 421 TITVEYORPNERTANLQSFMTNIDTKHNTVEOTIYINPLRYSAKETNVIISNGDEGST 480
QY 481 IIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTNDDYAQLGNNNDVINFGNIDSPYI 540
DB 481 IIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTNDDYAQLGNNNDVINFGNIDSPYI 540
QY 541 IKVISKYDPNKDDYTTTQQVTWQTINEYTGEPRTASYNNTAFSTSSQGGQGLDLPPEK 600
DB 541 IKVISKYDPNKDDYTTTQQVTWQTINEYTGEPRTASYNNTAFSTSSQGGQGLDLPPEK 600
QY 601 TYKIGDVWEDVDKDGQNTNDNEKPLSNVLVTLTYPDGSKSVRTDECKYQFDGLKNG 660
DB 601 TYKIGDVWEDVDKDGQNTNDNEKPLSNVLVTLTYPDGSKSVRTDECKYQFDGLKNG 660
QY 661 LTYKITPETPEGYTPTLKHSGTNPALDSEGNVSWVTINGQDDMTIDSGFYQTPKYSIGNY 720
DB 661 LTYKITPETPEGYTPTLKHSGTNPALDSEGNVSWVTINGQDDMTIDSGFYQTPKYSIGNY 720
QY 721 VWYDTNKGIGQDDEKIGISGVKVTLDKENGNIISTTTDENGKYQFDNLNSGNYIVHFDK 780

DB 721 VWYDTNKGIGQDDEKIGISGVKVTLDKENGNIISTTTDENGKYQFDNLNSGNYIVHFDK 780
QY 781 PSQWTTQTTTDSGDDDEQDADGSEVHVITTDHDDFSIDNGYDDDESDDSDSDSDSD 840
DB 781 PSQWTTQTTTDSGDDDEQDADGSEVHVITTDHDDFSIDNGYDDDESDDSDSDSDSD 840
QY 841 SD 900
DB 841 SD 900
QY 901 SD 960
DB 901 SD 960
QY 961 SD 1020
DB 961 SD 1020
QY 1021 SD 1080
DB 1021 SD 1080
QY 1081 LKRRKRNKKN 1092
DB 1081 LKRRKRNKKN 1092
RESULT 2
ABM79019
ID ABM79019 standard; protein; 1092 AA.
XX
XX ABM79019;
XX AC AC
XX XX
DT 15-JAN-2004 (first entry)
XX
DE Staphylococcus epidermidis polypeptide.
XX
XX Infection; antibacterial; vaccine.
XX
XX Staphylococcus epidermidis.
XX
XX WO2003076470-A1.
XX
XX 19-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-US006415.
XX
XX 05-MAR-2002; 2002US-0361324P.
XX
XX (INHI-) INHIBITEX INC.
XX (TEXA) UNIV TEXAS A & M SYSTEM.
XX
XX Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
XX Robbins J, Vernachio J, Bowden MG;
XX WPI; 2003-722324/68.
XX
XX
XX New antibody recognizing a Staphylococcus epidermidis protein comprising
XX Sdrg N1X2N3, Sdrg N2N3 or Sdrg2 useful for preparing a composition for
XX treating or preventing a coagulase-negative Staphylococcus infection.
XX
XX Claim 16; Page 36-37; 78pp; English.
XX
XX The present sequence comprises the protein sequence of a polypeptide of a
XX coagulase-negative Staphylococcus epidermidis. A claimed monoclonal
XX antibody recognises this protein and is used in a claimed method of
XX treating or preventing a coagulase-negative staphylococcal infection in a
XX human or animal, e.g. a nosocomial coagulase-negative staphylococcal
XX infection in low birth weight infants
XX
XX Sequence 1092 AA;

[illegible]

```
QY 184 EDSHVSDFPANSKIKESNTESKENTIEQPNKVKEDSTTSQPSGYTNIIDEKISNQDELIN 243
Db 181 ENSRVSDFPANSKIKESNTESKENTIEQPNKVKEDSTTSQPSGYTNIIDEKISNQDELIN 240
QY 244 LPINEYENKARPLSTSAQPSKRVNQLAAEQGSNNVHLIKVTDQSIPEGYDSDGVI 303
Db 241 LPINEYENKARPLSTSAQPSKRVNQLAAEQGSNNVHLIKVTDQSIPEGYDSDGII 300
QY 304 KAHAENLIYDVTEVDDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGELIATGT 363
Db 301 KAHAENLIYDVTEVDDKVKSGDTMTVDIDKNTVPSDLTDSFAPIKIKNSGELIATGT 360
QY 364 YDNKNKQITTYTFTDYVVKYENIKAHKLKLTYSIDSKVPNNNTKLDVYKTSALSVNKTIT 423
Db 361 YDNTNKQITTYTFTDYVVKYENIKAHKLKLTYSIDSKVPNNNTKLDVYKTSALSVNKTIT 420
QY 424 VEYORPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDEGSIID 483
Db 421 VEYORPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDEGSIID 480
QY 484 DSTIIVKVKYKGDGNQNLPSNRIDYSEYEDVTNDYDQAGLNNDVNIIFGNIDSPYIIVK 543
Db 481 DSTIIVKVKYKGDGNQNLPSNRIDYSEYEDVTNDYDQAGLNNDVNIIFGNIDSPYIIVK 540
QY 544 ISKYDPNKDDYTTTQQVTVMQTTINEYTGEBRTASYDNTIAFSTSSGQGGDLPPPEKTYK 603
Db 541 ISKYDPNKDDYTTTQQVTVMQTTINEYTGEBRTASYDNTIAFSTSSGQGGDLPPPEKTYK 600
QY 604 IGDVWVEDVDKQIGONTNDNEKPLSNVLTITVDPGTSKSVRTDEGKYQFDGLKGLTY 663
Db 601 IGDVWVEDVDKQIGONTNDNEKPLSNVLTITVDPGTSKSVRTDEGKYQFDGLKGLTY 660
QY 664 KITPETPEGYTPTLKHSNTPALDSEGNVSVMTINGQDDMTIDSGFTQTPKYSIGNYVWY 723
Db 661 KITPETPEGYTPTLKHSNTPALDSEGNVSVMTINGQDDMTIDSGFTQTPKYSIGNYVWY 720
QY 724 DTNKGIGQDDEKIGISGVKVTLDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSG 783
Db 721 DTNKGIGQDDEKIGISGVKVTLDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSG 780
QY 784 MTQTTDSGDDDEQADAGEVHVHTITDHDDFSINDGYDDSDSDSDSDSDSDSDSDSDS 843
Db 781 MTQTTDSGDDDEQADAGEVHVHTITDHDDFSINDGYDDSDSDSDSDSDSDSDSDSDS 833
QY 844 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 903
Db 834 ----- 833
QY 904 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 963
Db 834 ----- 833
QY 964 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1023
Db 834 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 880
QY 1024 GSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1083
Db 881 -----NSSDKNTKDLPTDGTGANEDHDKGTLGALLG 921
QY 1084 RKNRKNKN 1092
Db 922 RKNRKNKN 930
RESULT 4
AAY83171
ID AAY83171 standard; protein; 991 AA.
XX
AC AAY83171;
XX
DT 24-JUL-2000 (first entry)
XX
```

```
DE Cell wall protein SdrG.
XX SdrF; SdrG; SdrH; coagulase negative; staphylococcus; scepticemia;
KW osteomyelitis; endocarditis; immune response; vaccine; graft; stent;
KW intravenous catheter; heart valve; cardiac.
XX Staphylococcus sp.
XX Key Location/Qualifiers
PH Misc-difference 14
FT Misc-difference /note= "Position encoded by TAG stop codon"
FT Misc-difference 33
FT Misc-difference /note= "Position encoded by TGA stop codon"
FT Misc-difference 964
FT Misc-difference /note= "Position encoded by TAA stop codon"
FT Misc-difference 980
FT Misc-difference /note= "Position encoded by TAG stop codon"
FT Misc-difference 989
FT Misc-difference /note= "Position encoded by TAA stop codon"
XX WO200012689-A1.
XX 09-MAR-2000.
XX 31-AUG-1999; 99WO-US019728.
XX 31-AUG-1998; 98US-0098443P.
PR 25-JAN-1999; 99US-0117119P.
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
PA (TEXA ) UNIV TEXAS A & M SYSTEM.
XX Poster TJ, Hook M, Davis S, Hartford O, McCrear K, Ni Eidhin D;
XX WPI; 2000-256637/22.
DR N-PSDB; AA293534.
XX Recombinant or synthetic proteins from coagulase-negative staphylococci
PT useful for prevention, treatment and diagnosis of staphylococcal
PT infections bind soluble and immobilized fibrinogen.
XX Claim 8; Fig 3; 104pp; English.
XX Isolated Staphylococcus Sdr cell wall proteins which bind both soluble
CC and immobilized fibrinogen are useful for treating or preventing
CC coagulase-negative staphylococcal infection such as scepticemia,
CC osteomyelitis or endocarditis, and for inducing immune responses in
CC patients. The cell wall proteins are also useful for reducing coagulase-
CC negative staphylococci infection of indwelling medical devices such as
CC vascular grafts, vascular stents, intravenous catheters, artificial heart
CC valves and cardiac assist devices. The cell wall associated proteins are
CC able to inhibit staphylococcal adhesion to immobilised extracellular
CC matrix or host cells present on the surface of implanted biomaterials
XX Sequence 991 AA;
SQ
```

```
Query Match 79.4%; Score 4485; DB 3; Length 991;
Best Local Similarity 81.1%; Pred. No. 3.7e-210;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;
QY 4 KKNLLTKKKKIPANKSNKYAIRKFTVGTASIVIGATILFGLGHNEAKAENSVDQVKDSN 63
Db 35 KKNLLTKKKKIPANKSNKYAIRKFTVGTASIVIGATILFGLGHNEAKAENSVDQVKDSN 94
QY 64 TDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 123
Db 95 MDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 153
QY 124 VDENEATFLQKTPQDNTLHTEEVKESSESSNIDTAQQPSHTTINREESVQTSNV 183
Db 154 VDENEATFLQKTPQDNTLHTEEVKESSESSNIDTAQQPSHTTINREESVQTSNV 213
QY 184 EDSHVSDFPANSKIKESNTESKENTIEQPNKVKEDSTTSQPSGYTNIIDEKISNQDELIN 243
```

Db 214 ENRSVDFANSKIIIESNTESKENTIEQNKVREDSITSQPSYKNIDEKISNQDELLN 273
Qy 244 LPINEYENKARPLSTTSAQSIKRVTVNQLAARQGSNVNHLIKVTDQSITEGYDDSEGI 303
Db 274 LPINEYENKVRPLSTTSAQSPSKRVTVNQLAARQGSNVNHLIKVTDQSITEGYDDSGII 333
Qy 304 KAHAENLIYDVTFEVDKVKSGDTMTVDIKNTVPSDLTDSFTIPIKIDNSGEIITATGT 363
Db 334 KAHAENLIYDVTFEVDKVKSGDTMTVDIKNTVPSDLTDSFAIPKIDNSGEIITATGT 393
Qy 364 YDNKNKQITVTFDYDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTLSSVKNKTI 423
Db 394 YDNKNKQITVTFDYDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTLSSVKNKTI 453
Qy 424 VEYORPENRNTANLQSMFTNIDTKNHTVEQIYINPLYSKAKETNVIISGNGEGSTIID 483
Db 454 VEYORPENRNTANLQSMFTNIDTKNHTVEQIYINPLYSKAKETNVIISGNGEGSTIID 513
Qy 484 DSTIIVKVKGNQNLPSNRNRIYDYSEYEDVTNDDYAQLGNRNDVNFNIDSPYIIKV 543
Db 514 DSTIIVKVKGNQNLPSNRNRIYDYSEYEDVTNDDYAQLGNRNDVNFNIDSPYIIKV 573
Qy 544 ISKYDKNKDYTTIQCTVTMOTTTINEYTGERTASVDNTIAPSTSGQGGDLPPKTYK 603
Db 574 ISKYDKNKDYTTIQCTVTMOTTTINEYTGERTASVDNTIAPSTSGQGGDLPPKTYK 633
Qy 604 IGVYVWEDVDKQIQNTDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKNGLTY 663
Db 634 IGVYVWEDVDKQIQNTDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKNGLTY 693
Qy 664 KITFETPEGYPTPLKHSQTNPALDSGNSVWVTINGQDDMTIISGFGYQTPKYSLGNYVY 723
Db 694 KITFETPEGYPTPLKHSQTNPALDSGNSVWVTINGQDDMTIISGFGYQTPKYSLGNYVY 753
Qy 724 DTNKGQIGDDEKIGSKVYTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFKPSSG 783
Db 754 DTNKGQIGDDEKIGSKVYTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFKPSSG 813
Qy 784 MTQTTTSGDDQDADGEEVHVITTDHDDFSIDNGYDDSDSDSDSDSDSDSDSDSDSDS 843
Db 814 MTQTTTSGDDQDADGEEVHVITTDHDDFSIDNGYDDSDSDSDSDSDSDSDSDSDSDS 866
Qy 844 DS 903
Db 867 ----- 866
Qy 904 DS 963
Db 867 ----- 866
Qy 964 DS 1023
Db 867 DS 913
Qy 1024 GSDS 1083
Db 914 -----NSSDKNTKDKLPDTCANEDHSDSGTLLGALPAGLILGK 954
Qy 1084 RKXNRKNK 1092
Db 955 RKXNRKNK 963

RESULT 5
AA70120
ID AA70120 standard; protein; 991 AA.
XX AC
XX AA70120;
DT 06-JUN-2000 (first entry)
XX DE Staph. epidermidis serine-aspartate repeat region protein SdrG.

XX Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;
KW microbial surface components recognising adhesive matrix molecules;
KW collagen binding protein; CBP; CNA; fibrinogen binding protein;
KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;
KW fibronectin binding protein; Staphylococcus infection;
KW serine-aspartate repeat region protein; SDR protein; SdrG.
XX Staphylococcus epidermidis.
OS
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 14 /note= "Encoded by in-frame stop codon TAG"
FT Misc-difference 33 /note= "Encoded by in-frame stop codon TGA"
FT Misc-difference 964 /note= "Encoded by in-frame stop codon TAA"
FT Misc-difference 980 /note= "Encoded by in-frame stop codon TAG"
FT Misc-difference 989 /note= "Encoded by in-frame stop codon TAA"
PN WO200012131-A1.
XX 09-MAR-2000.
XX 31-AUG-1999; 99WO-US019727.
XX 31-AUG-1998; 98US-0098439P.
PA (INHI-) INHIBITEX INC.
PA (TEXA) UNIV TEXAS A & M SYSTEM.
PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
PI Patti JM, Foster TJ, Hook M;
XX WPI: 2000-237781/20.
DR N-PSDB; AA251202.
XX
PT Composition used for generating immune response or for inhibiting
PT microbial colonization in an animal comprises antibodies that bind
PT collagen binding protein, fibrinogen binding protein and, optionally,
PT fibronectin binding protein.
XX
PS Claim 8; Fig 4; 115pp; English.
XX
CC The patent discloses multicomponent vaccines containing selected
CC combinations of bacterial binding proteins termed MSCRAMM (microbial
CC surface components recognising adhesive matrix molecules) or their
CC antibodies. A vaccine composition is provided that includes collagen
CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein
CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB), and
CC optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are
CC useful for imparting protection against a broad spectrum of
CC staphylococcal strains and for inhibiting microbial colonisation.
CC especially of Staphylococcus aureus, in an animal. The combinations can
CC also be used to select donor blood pools for the preparation of purified
CC blood products for passive immunisation. The present sequence is a serine
CC -aspartate repeat region protein, SdrG from Staphylococcus epidermidis.
CC The Sdr protein is useful in vaccine preparation in combination with
CC specific bacterial binding proteins. These vaccines can be used to treat
CC a broad spectrum of bacterial infections, including those arising from
CC both coagulase-positive and coagulase-negative bacteria
XX
SQ Sequence 991 AA;
XX
XX Query Match 79.4%; Score 4485; DB 3; Length 991;
XX Best Local Similarity 81.1%; Pred. No. 3.7e-210;
XX Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;
Qy 4 KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVKDSN 63
Db 35 KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVKDSN 94

SQ		Sequence 892 AA;	
Query Match		76.3%; Score 4307; DB 6; Length 892;	
Best Local Similarity		80.4%; Pred. No. 1.5e-201;	
Matches		846; Conservative 20; Mismatches 26; Indels 160; Gaps 3;	
Qy	41	LFLGLHNEAKAENSVDVKDSNTDDELSDSDSDDEKNDVINNNQSDTDDNNQI1K 100	
Db	1	MFGLGHNEAKAENTVDVKDSNMDDSDSDSDNEEKNDVINNSQINTDDNQ-IK 59	
Qy	101	KEETNYDGIKEKSDTESTTNDVNEATFLQKTPQDNTLHTEEVKSSSVESNSSI 160	
Db	60	KEETNSDAIENRSDITSTTNDVNEATFLQKTPQDNTLHTEEVKSSSVESNSSM 119	
Qy	161	DTAQPSHTTINREESVQTSNDVEDSHVSDPANSKIKESNTSGKEENTIEQPNKYKPS 220	
Db	120	DTAQPSHTTINSEASIQTSNEENSKVSDPANSKIKESNTSGKEENTIEQPNKYKPS 179	
Qy	221	TTSQPSGYTNIDEKISNQDELLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQSN 280	
Db	180	ITSQPSGYKNIDEKISNQDELLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQSN 239	
Qy	281	VNHLIKVTQSIQTEGVDDSGEVIKAHDAENLIYDVTFEVDDKVKSGDTMTVDIKNTVPS 340	
Db	240	VNHLIKVTQSIQTEGVDDSGEVIKAHDAENLIYDVTFEVDDKVKSGDTMTVDIKNTVPS 299	
Qy	341	DLTDSFTPIKIDNSGEIITATGYDNKNKQITVFTDYDVKYENIKAKHLKLTYSIDKSKV 400	
Db	300	DLTDSFAPKIDNSGEIITATGYDNKNKQITVFTDYDVKYENIKAKHLKLTYSIDKSKV 359	
Qy	401	PNNNTKLDVEYKLTALSSVAKTIIVEQPNENRNTANLQSMFTNIDTKNHTVEQTIYNPL 460	
Db	360	PNNNTKLDVEYKLTALSSVAKTIIVEYQKPNENRNTANLQSMFTNIDTKNHTVEQTIYNPL 419	
Qy	461	RYSAKETNVIISNGDEGFTIIDSTIIKVKYVGDQNLPSNRIYDYSEYEDVTNDVYA 520	
Db	420	RYSAKETNVIISNGDEGFTIIDSTIIKVKYVGDQNLPSNRIYDYSEYEDVTNDVYA 479	
Qy	521	QLGNNDVNIINFGNIDSPYIIKISKYDPNKODYTTIQOVTVMQTTINEYTGFEFTASYD 580	
Db	480	QLGNNDVNIINFGNIDSPYIIKISKYDPNKODYTTIQOVTVMQTTINEYTGFEFTASYD 539	
Qy	581	NTIAPSTSSGQGGDLPPKTYKIGYVWEDVDKQIQNTNDNEKPLSNVLTLPDGT 640	
Db	540	NTIAPSTSSGQGGDLPPKTYKIGYVWEDVDKQIQNTNDNEKPLSNVLTLPDGT 599	
Qy	641	SKSVRTDEGKYQFDGLKNGLTVKITFETPEGYTPTLKHSGTNPALDSGNSVWVTINGQ 700	
Db	600	SKSVRTDEGKYQFDGLKNGLTVKITFETPEGYTPTLKHSGTNPALDSGNSVWVTINGQ 659	
Qy	701	DMTIDSGFYQTPKYSILGNVYVYDTNKGIGQDDEKIGSVKVTLDENGNIIISTTTTDS 760	
Db	660	DMTIDSGFYQTPKYSILGNVYVYDTNKGIGQDDEKIGSVKVTLDENGNIIISTTTTDS 719	
Qy	761	NGKYQFDNLNSGNYIYHFKPQSGMTQTTTDSGDDDEQDADGERVHTTITDHDPSIDNGY 820	
Db	720	NGKYQFDNLNSGNYIYHFKPQSGMTQTTTDSGDDDEQDADGERVHTTITDHDPSIDNGY 779	
Qy	821	YDDESD 880	
Db	780	YDDESD 795	
Qy	881	SD 940	
Db	796	SD 795	
Qy	941	SD 1000	
Db	796	SD 832	
Qy	1001	SD 1060	
Db	833	SD 860	

Qy	1061	DVSGSGTLLGTLFAGLGALLLGRKRKRKRKN 1092	
Db	861	DHDSKGTLLGALFAGLGALLLGRKRKRKN 892	
RESULT 7			
ABU42520			
ID	ABU42520	standard; protein; 670 AA.	
XX			
AC	ABU42520;		
DT	19-JUN-2003	(first entry)	
DE		Protein encoded by Prokaryotic essential gene #28047.	
XX			
KW		Antisense; prokaryotic essential gene; cell proliferation; drug design.	
OS		Staphylococcus epidermidis.	
XX			
PN	WO200277183-A2.		
XX			
PD	03-OCT-2002.		
XX			
PF	21-MAR-2002; 2002WO-US009107.		
XX			
PR	21-MAR-2001; 2001US-00815242.		
PR	06-SEP-2001; 2001US-00948993.		
PR	25-OCT-2001; 2001US-0342923P.		
PR	08-FEB-2002; 2002US-00072851.		
PR	06-MAR-2002; 2002US-0362699P.		
XX			
PA	(ELIT-) ELITRA PHARM INC.		
XX			
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;		
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;		
XX			
DR	WPI; 2003-029926/02.		
XX	N-PSDB; ACA46390.		
PT		New antisense nucleic acids, useful for identifying proteins or screening	
PT		for homologous nucleic acids required for cellular proliferation to	
PT		isolate candidate molecules for rational drug discovery programs.	
XX			
PS	Claim 25; SEQ ID NO 70444; 1766pp; English.		
XX			
CC		The invention relates to an isolated nucleic acid comprising any one of	
CC		the 6213 antisense sequences given in the specification where expression	
CC		of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC		(1) a vector comprising a promoter operably linked to the nucleic acid	
CC		encoding a polypeptide whose expression is inhibited by the antisense	
CC		nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC		polypeptide or its fragment whose expression is inhibited by the	
CC		antisense nucleic acid; (4) an antibody capable of specifically binding	
CC		the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
CC		proliferation or the activity of a gene in an operon required for	
CC		proliferation; (7) identifying a compound that influences the activity of	
CC		the gene product or that has an activity against a biological pathway	
CC		required for proliferation, or that inhibits cellular proliferation; (8)	
CC		identifying a gene required for cellular proliferation or the biological	
CC		pathway in which a proliferation-required gene or its gene product lies	
CC		or a gene on which the test compound that inhibits proliferation of an	
CC		organism acts; (9) manufacturing an antibiotic; (10) profiling a	
CC		compound's activity; (11) a culture comprising strains in which the gene	
CC		product is overexpressed or underexpressed; (12) determining the extent	
CC		to which each of the strains is present in a culture or collection of	
CC		strains; or (13) identifying the target of a compound that inhibits the	
CC		proliferation of an organism. The antisense nucleic acids are useful for	
CC		identifying proteins or screening for homologous nucleic acids required	
CC		for cellular proliferation to isolate candidate molecules for rational	
CC		drug discovery programs, or for screening homologous nucleic acids	
CC		required for proliferation in cells other than S. aureus, S. typhimurium,	
CC		K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of	

CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 670 AA;

Query Match 59.4%; Score 3351.5; DB 6; Length 670;
Best Local Similarity 96.3%; Pred. No. 3.4e-155;
Matches 648; Conservative 0; Mismatches 2; Indels 23; Gaps 2;
QY 440 MFTNIDTKHVEQTIYINPLRYSAKETNWNISGDEGSTIIDSTIIKVKYKVDNQNL 499
Db 1 MFTNIDTKHVEQTIYINPLRYSAKETNWNISGDEGSTIIDSTIIKVKYKVDNQNL 60
QY 500 PDSNRIYDYSEYEDVNDVAQLGNNDVNFNFGNIDSPYIIKVKYKVDNQDYYTIIQ 559
Db 61 PDSNRIYDYSEYEDVNDVAQLGNNDVNFNFGNIDSPYIIKVKYKVDNQDYYTIIQ 120
QY 560 TVTMTTINEYTGFRFASVNDTIAPSTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 619
Db 121 TVTMTTINEYTFE---ASYDNTIAFSTSSGQGGDLPPKTYKIGYVWEDVDKGIQN 177
QY 620 TNDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKGLTYKTIPTPEGYTPTLK 679
Db 178 TNDNEKPLSNVLTLYPDGTSKSVRTDEGKYQFDGLKGLTYKTIPTPEGYTPTLK 237
QY 680 SGTNPALDSGNSVWVTVINGQDDMTIDSFGYQTPKYSGLGNVWYDNTNKGIGQDDEKIS 739
Db 238 SGTNPALDSGNSVWVTVINGQDDMTIDSFGYQTPKYSGLGNVWYDNTNKGIGQDDEKIS 297
QY 740 GVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQDA 799
Db 298 GVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQDA 357
QY 800 DGEVHVITTDHDDFSIDNGYVDE-----SDSDSDSDSDSDS 839
Db 358 DGEVHVITTDHDDFSIDNGYVDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 417
QY 840 DS 899
Db 418 DS 477
QY 900 DS 959
Db 478 DS 537
QY 960 DS 1019
Db 538 DS 597
QY 1020 DS 1079
Db 598 DS 657
QY 1080 LLGKRRKRNKN 1092
Db 658 LLGKRRKRNKN 670

RESULT 8
ABM79020
ID ABM79020 standard; protein; 549 AA.
XX
AC ABM79020;
XX
XX
DT 15-JAN-2004 (first entry)
XX
DE Staphylococcus epidermidis polypeptide.
XX
KW Infection; antibacterial; vaccine.
XX
OS Staphylococcus epidermidis.

XX WO2003076470-A1.
PN 18-SEP-2003.
XX 05-MAR-2003; 2003WO-US006415.
XX 05-MAR-2002; 2002US-0361324P.
XX (INH1-) INHIBITEX INC.
PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
PI Robbins J, Vernachio J, Bowden MG;
XX WPI; 2003-722324/68.
XX New antibody recognizing a Staphylococcus epidermidis protein comprising
PT SdrG N1N2N3, SdrG N2N3 or SdrGTR2 useful for preparing a composition for
PT treating or preventing a coagulase-negative Staphylococcal infection.
XX Claim 20; Page 37; 78pp; English.
XX The present sequence comprises the protein sequence of a polypeptide of a
CC coagulase-negative Staphylococcus epidermidis. A claimed monoclonal
CC antibody recognises this protein and is used in a claimed method of
CC treating or preventing a coagulase-negative staphylococcal infection in a
CC human or animal, e.g. a nosocomial coagulase-negative staphylococcal
CC infection in low birth weight infants
XX Sequence 549 AA;

Query Match 49.4%; Score 2787; DB 7; Length 549;
Best Local Similarity 98.7%; Pred. No. 7.9e-128;
Matches 542; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 52 EENSVDQVKSDNTDDELSDS 111
Db 1 EENSVDQVKSDNTDDELSDS 60
QY 112 KESEORTSTTNDVNEATFLQKTPQDNTHLTETEEVKESSESSVESNSSIDTAAQPSHTTI 171
Db 61 KESEORTSTTNDVNEATFLQKTPQDNTHLTETEEVKESSESSVESNSSIDTAAQPSHTTI 120
QY 172 NREESVQTSNDVEDSHVSDPFANSKIKESNTESGKEENTIEQPNKVKESSTTSQPSGYTNI 231
Db 121 NREESVQTSNDVEDSHVSDPFANSKIKESNTESGKEENTIEQPNKVKESSTTSQPSGYTNI 180
QY 232 DEKISNQDELMLPINEYENKARPLSTTSAQPSIKRVTNQLAARQGSNNVNLIKVTDQS 291
Db 181 DEKISNQDELMLPINEYENKARPLSTTSAQPSIKRVTNQLAARQGSNNVNLIKVTDQS 240
QY 292 ITEGYDDSEGVTKAHDENLIYDVTFFVDDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKI 351
Db 241 ITEGYDDSEGVTKAHDENLIYDVTFFVDDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKI 300
QY 352 KNSGEIATGYDNKNKQITVTFDYDYDKYENIKAHKLTSYIDKSKVPNNNTKLDVEY 411
Db 301 KNSGEIATGYDNKNKQITVTFDYDYDKYENIKAHKLTSYIDKSKVPNNNTKLDVEY 360
QY 412 KTALESVNKTIITVEYQRPENRNTANLQSMFTNIDTKHVEQTIYINPLRYSAKETNVI 471
Db 361 KTALESVNKTIITVEYQRPENRNTANLQSMFTNIDTKHVEQTIYINPLRYSAKETNVI 420
QY 472 SNGDDEGSTIIDSTIIKVKYKVDNQNLPSNRIYDYSEYEDVNDVAQLGNNDVNFN 531
Db 421 SNGDDEGSTIIDSTIIKVKYKVDNQNLPSNRIYDYSEYEDVNDVAQLGNNDVNFN 480
QY 532 FGNIDSPYIIKVKYKVDNQDYYTIIQOTVTMTTINEYTGFRFASVNDTIAPSTSSGQ 591
Db 481 FGNIDSPYIIKVKYKVDNQDYYTIIQOTVTMTTINEYTGFRFASVNDTIAPSTSSGQ 540
QY 592 QGGDLPPPEK 600

Db 541 GQGLPPEK 549

RESULT 9

AAV08643

ID AAY08643 standard; protein; 1166 AA.

AC AAY08643;

XX 20-MAR-2003 (revised)

DT 09-AUG-1999 (first entry)

XX XX

DE S. aureus SdrE protein.

XX Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;

KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;

KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;

KW extracellular matrix; vascular graft; vascular stent; vaccine;

KW intravenous catheter; artificial heart valve; cardiac assist device;

KW antibacterial.

XX Staphylococcus aureus.

OS Staphylococcus aureus.

XX W09927109-A2.

FN 03-JUN-1999.

XX 25-NOV-1998; 98WO-US025246.

PF 26-NOV-1997; 97US-0066815P.

PR 31-AUG-1998; 98US-0098427P.

XX (INH-) INHIBITEX INC.

PA (FORP-) FORP T/A BIORESEARCH IRELAND.

PA (TEXA) UNIV TEXAS A & M.

PA (PATT) PATTI J M.

PA (JOSE) JOSEFFSON E.

PA (EIDH) EIDH D N.

PA (HOOK) HOOK M A O.

PA (PERK) PERKINS S E.

XX Patti JM, Foster TJ, Joseffson E, Eidhin DN, Hook MAO;

PI Perkins SE;

XX WPI: 1999-357844/30.

DR N-PSDB; AAX77594.

XX Staphylococcus aureus fibrinogen-binding proteins for treating

PT septicemia, osteomyelitis, mastitis or endocarditis.

XX Claim 8; Fig 9; 143pp; English.

XX This invention describes novel Staphylococcus aureus fibrinogen-binding

CC proteins that bind both the alpha and beta fibrinogen chains. The

CC proteins (and their encoding nucleic acids) are ClfB, SdrC, SdrD and

CC SdrE. Staphylococcus aureus is thought to utilize fibrinogen to adhere

CC to medical devices, binding proteins that bind both the alpha and beta

CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as

CC competitive inhibitors to block this binding. Antibodies against ClfB,

CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.

CC The proteins of the invention can be used in a pharmaceutical composition

CC for the treatment of Staphylococcus aureus infection e.g. septicemia,

CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of S.

CC aureus to the extracellular matrix. The proteins or their fragments may

CC be used to coat a medical device to reduce the S. aureus infection of an

CC indwelling medical device, especially where the medical device is

CC selected from the group consisting of vascular grafts, vascular stents,

CC intravenous catheters, artificial heart valves, and cardiac assist

CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or

CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat

CC region or a gene encoding it may be used as an identifying probe for the

CC

CC identification of genes and encoding proteins from Staphylococcus aureus

CC (other than Cifa), S. hemolyticus, S. lugdenensis, and S. schleriferi.

CC The proteins of the invention have antibacterial activity. (Updated on 20

CC -MAR-2003 to correct PA field.)

XX

SQ Sequence 1166 AA;

Query Match 47.8%; Score 2698.5; DB 2; Length 1166;

Best Local Similarity 48.2%; Pred. No. 4.1e-123;

Matches 595; Conservative 158; Mismatches 267; Indels 215; Gaps 26;

QY 1 MINKKN-LLTKKKPIANKSNKYARKFTVGTASIVIGATLLPGLGHNKAKAENSVDV 59

DB 1 MINRDNKKAITKKGMISNRLNFKSIRKTYVTGASILVLTFLPGLNQAKAAENT--ST 58

QY 60 KDSNTDDELSDNDSSDEEKNDVINNQSINTDDNNQIIKKEETNNYDGIKRSERDTE 119

DB 59 ENAKQDDATSDNKEVSENNSTENNSTNP-----IKKE--INTSQPEAKKETS 110

QY 120 STTNVDENEATFLQKTPQDNTHLTBEVEKES---SSVESNSISIDTAQPSHTTIRES 176

DB 111 SSTQKQNNVTATTETKPN--IEKENVPSTDKTATEDSVILEKCAPNNT--NNDVT 166

QY 177 VQTSNVEDSHVDFANSKIKESNT---ESGKEENTIEQPNKVKEDSTTSQPSGYTNIDE 233

DB 167 TKPS-----TSEPSTSEIQTKPTTPQBSTNIENSQPTPSKVD---NQVTDATNPKE 216

QY 234 KIS-NODELLNLP-----INYEENKARPLSTTSAQPSIKRV-----TYNQLAEEQ 277

DB 217 PVNVSKELKNPEKELVNRDNTDHTKPVATAPTSAVKRVNAKRFVAQAAPAAVA 276

QY 278 GSNVHLIKVTDQSIGTEGYDDSEGVKAHAENLIYDVTFEVDVKVSGDTMTVDIDKNT 337

DB 277 SNNVNDLIKVTQTIKVG-DGKNVAAAHGDKIEYDTEFTIDNKKVKGDTMTINDKVN 335

QY 338 VPSDLTDSSTIPIKIDNSGEIATGTYNKNQIYTTFTDYDKYENIKAHKLTSLVIDK 397

DB 336 IPSDLTDRKNDPIDITDPSGEVIAGTFDKATQITFTDYVDYDKIEDIKSLRLTLYSIDK 395

QY 398 SKVPNNNTKLDVEYKTALSSVNKTITVEYQRPNNENRNTANLQSMFTNIDTKNHTVEQTIYI 457

DB 396 KTVP-NETSLNLTFFATAGKETSQNVTVQDPVHVGDSNIQSIFFTKLDEKDKQIEQIYV 454

QY 458 NPLRYSAKETNVNISNG-----DEGSTIIDSTIIKVKYKGDGNQNLPSNRIYD 508

DB 455 NPLKKSATNTKVDIAGSQVDDYGNIKLGNSTIIDQNTBIKVKVNSDQOLPQSNRIYDF 514

QY 509 SEYEDVTND-DYAQLGNNDNVNFGNIDSPYIIKVISKYDKNKDDYTTIQQVTTWQTTI 567

DB 515 SQYEDVTSQDNKKSFSNNVATLDFGDIINSAYIIKVVSKYTPSTDSGELDIAQGTSMRTT- 573

QY 568 NEYTGFEFTASYDNTIAFSTSSQGGQGLD-PPEKTYKIGDYVWEDVDKQIQNTNNEKP 626

DB 574 DKY-GYNTAGSYNFIVTSNDTGGDGTVKPEKLYKIGDYVWEDVDKGVQGTDSKEP 632

QY 627 LSNVLTLTYPDGTSKSVRTDEBGKQFQPLGNGLYKITFTEPEGYTPTFLKHSIGNPAL 686

DB 633 MANVLTLTYPDGTTKSVRTDANGHYEFGGLKDGETYTVKFTPTGYLPTKVGNTGDGEK 692

QY 687 DSEGNVWVTINGQDDMTIDSGFYQTPKYSLGNVYVYDNTKQGIQGDDEKGISGVKVTLK 746

DB 693 DSGSSVTVKINGKODMSLDTFGYPEKYNLGDYVWEDNKGQIQDANEPEGIKDVKVTLK 752

QY 747 DENGNIIISITT----- 757

DB 753 DSTGKVIQITTTDASGKYFTDLNNGNYTFETPAGYTPVKNYTTADDKDSNGLTTGV 812

QY 758 ----- 757

DB 813 IKDADNMTLDRGPKTPKYSGLGVWYVNSKQKQDSTEKGIKDVTVTTLQNEKGEVIGTT 872

QY 758 -TDENKIQFDNLNSGNYIVHEDKPGMTQTITDSDGDDDEQDADGGEVHVHTIHDHDFSI 816

Db 816 ADMWTLDRGYKTPKYSGLDYVWDSNKGQKQSTEGIKVDVTVTLQNEKEVIGTKTD 875
 Qy 760 ENGKYQDNLNSGNYIVHFKPSGWTQTTTDSGDDDEQDADGEEVHVTTITDHDQDFSIDNG 819
 Db 876 ENGKYRFDNLDSGKYKVIPEKPAQLTQVNTTDEDD-KDADGGEVDVTTITDHDQDFTLNG 934
 Qy 820 YDDESDS 879
 Db 935 YFEEDT-----SDS 983
 Qy 880 DS 939
 Db 984 DS 1043
 Qy 940 DS 999
 Db 1044 DS 1103
 Qy 1000 DS 1058
 Db 1104 DSDAGKHTPVKPMW-----TTKDHNNKAKALPETGS 1134
 Qy 1059 NEDYGSKGTILGTLFAGLGALLL-GKRRKNRK 1089
 Db 1135 ENNGSNATLFGGLFAALGSLLLFGRRKQNK 1166
 RESULT 11
 ID ABM79015 standard; protein; 560 AA.
 AC ABM79015;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Staphylococcus epidermidis SdrG N1N2N3 domain.
 XX
 KW SdrG; surface protein; infection; antibacterial; vaccine.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN W02003076470-A1.
 XX
 PD 18-SEP-2003.
 XX
 PF 05-MAR-2003; 2003WO-US006415.
 XX
 PR 05-MAR-2002; 2002US-0361324P.
 XX
 PA (INH1-) INHIBITEX INC.
 XX (TEXA) UNIV TEXAS A & M SYSTEM.
 XX
 PI Patti JM, Hutchins JT, Hall A, Domanski P, Patel P, Hook M;
 PI Robbins J, Vernachio J, Bowden MG;
 XX
 DR WPI; 2003-722324/68.
 DR N-PSDB; ACF80624.
 XX
 XX New antibody recognizing a Staphylococcus epidermidis protein comprising
 PT SdrG N1N2N3, SdrG N2N3 or SdrGTR2 useful for preparing a composition for
 PT treating or preventing a coagulase-negative Staphylococcus infection.
 XX
 PS Claim 27; Page 24-25; 78pp; English.
 CC The present sequence comprises the protein sequence of the N1N2N3 region
 CC (amino acids 50-597), or putative A domain, of the SdrG surface protein
 CC of coagulase-negative Staphylococcus epidermidis. A claimed antibody
 CC recognises a protein selected from SdrG N1N2N3, N2N3 and TR2. The
 CC antibody may be a monoclonal antibody, including a chimeric, murine,
 CC humanized, human or single chain monoclonal antibody, which prevents a
 CC coagulase-negative staphylococcus infection in a human or animal by
 CC inhibiting binding of staphylococcus bacteria to fibrinogen. Such

CC antibodies can be used to treat or prevent staphylococcal infections
 CC including nosocomial coagulase-negative staphylococcal infections in low
 CC birth weight infants. A claimed vaccine comprises the isolated SdrG
 CC N1N2N3, N2N3 or TR2 protein
 XX
 SQ Sequence 560 AA;
 Query Match 46.5%; Score 2624.5; DB 7; Length 560;
 Best Local Similarity 92.3%; Pred. No. 6.6e-120;
 Matches 512; Conservative 17; Mismatches 25; Indels 1; Gaps 1;
 Qy 46 HNEAKAEENSVDKSDNTDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 105
 Db 7 HHHGSEENVQVDKSDNMDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 65
 Qy 106 NYDGIKRESDRSTETNVNDEATEFLQKTPQDNTLHTEEVKESSESSSSSIDTAQQ 165
 Db 66 SDAENRISKDITQSTTNVDENEATEFLQKTPQDNTLHTEEVKESSESSSSMDTAQQ 125
 Qy 166 PSHTTINRESVQTSNVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQP 225
 Db 126 PSHTTINSEASIQTSNDEENSRVSDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQP 185
 Qy 226 SGVTNIDEKISNDELLNLPINEYENKARPLSTTSAPSIKRVTVNQLAAEQGSNNVHLI 285
 Db 186 SSYKNIDEKISNDELLNLPINEYENKVRPLSTTSAPSKRVTVNQLAAEQGSNNVHLI 245
 Qy 286 KVTQDSITEGYDDSEGVKIKAHDAENLIYDVTPEVDDKVKSGDTMTVDIDKNTVPSDLTDS 345
 Db 246 KVTQDSITEGYDDSGIKAHDAENLIYDVTPEVDDKVKSGDTMTVDIDKNTVPSDLTDS 305
 Qy 346 FPIPKIKONGELIATGYDNKNKQIYTYTFTDYDKYENIKAHKLKLTSTYIDKSKVPPNNT 405
 Db 306 FAIPKIKDNGSEIATGYDNTNKKQIYTYTFTDYDKYENIKAHKLKLTSTYIDKSKVPPNNT 365
 Qy 406 KLDVEYKLTALSSVKNKTIITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLYSK 465
 Db 366 KLDVEYKLTALSSVKNKTIITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLYSK 425
 Qy 466 ETNNVNSGNGDEGTTIIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTNDDYQAQGN 525
 Db 426 ETNNVNSGNGDEGTTIIDSTIIKVKYKVDGNQNLPSNRIYDYSEYEDVTNDDYQAQGN 485
 Qy 526 NDVNFNGNIDSPYIIKVKSKYDPNKKDYTTTQOQVTMTQTTINEYTGEPRTASYDNTIAF 585
 Db 486 NDVNFNGNIDSPYIIKVKSKYDPNKKDYTTTQOQVTMTQTTINEYTGEPRTASYDNTIAF 545
 Qy 586 STSSGQGGDLPEK 600
 Db 546 STSSGQGGDLPEK 560
 RESULT 12
 ID ABU42327
 ID ABU42327 standard; protein; 1141 AA.
 AC ABU42327;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #27854.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN W0200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US0009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.

antibodies. A vaccine composition is provided that includes collagen binding protein or peptide, e.g. CNA, a fibrinogen binding protein preferably Clumping factor A (ClfA) or Clumping factor B (ClfB), and optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are useful for imparting protection against a broad spectrum of Staphylococcal strains and for inhibiting microbial colonisation, especially of *Staphylococcus aureus*, in an animal. The combinations can also be used to select donor blood pools for the preparation of purified blood products for passive immunisation. The present sequence is a serine-aspartate repeat region protein, SdrF from *Staphylococcus epidermidis*. The Sdr protein is useful in vaccine preparation in combination with specific bacterial binding proteins. These vaccines can be used to treat a broad spectrum of bacterial infections, including those arising from both coagulase-positive and coagulase-negative bacteria

Sequence 1802 AA:
SQ

Sequence 1802 AA;

Query Match 40.8%; Score 2303.5; DB 3; Length 1802;
Best Local Similarity 34.6%; Pred. No. 1.2e-103;
Matches 605; Conservative 146; Mismatches 313; Indels 685;

Qy	2	INKQNNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLPLGLGHNEAKAEEN-----S 55
Db	46	INKRVDEL-----SNKVNKYSIRKFTVGTASILVCGATLMPGAADNEAKAAEDNQLESAS 99
Qy	56	VQDVKOS--NTDDELS--DSNDQSSDEKNDVINN-----NOSIN 91
Db	100	KEBQKGRDNENSKLNQVDLONGSHSSEKTTNNVNNATEVKKVEAPTTSVSKPKANEAV 159
Qy	92	TDD-----NNQII---KKEETNNYDGIK-----RSEDRTESSTNNVDEN 127
Db	160	TNESTPKTTTEAPTVNEESAETPKTSTTQQDSTTEKNPNSLKONLNESSITTSKESKTDH 219
Qy	128	EATFLQ-KTQDNTHLTTEEVK-ESSSVESNSSIDTAQOPS-----167
Db	220	STKQAQMSITNSKLNLDTWDSPTQSEKTSQANNDSTDNQSAQSKLDSPKSPQKYVTKFN 279
Qy	168	-----HTTIN-REBSVOTSDNVEDSHVDFANSKIKESNTSESGKEENTIEQPNKVE 218
Db	280	DEPTQDVEHTTTLKTPSVSTDSVNDK--ODYTRSAV-----ASLGVDNSTETAITNAVR 333
Qy	219	DSTTSOPSGYTNIDEKI-----SNQDELLMLP--INEYENKARPLSTTSQAQSIKR 267
Db	334	DNLDLRAASREQINEAIIAEALKKDFSNPDYGVDTPLALNRSQKSNP--HKSASP--R 388
Qy	268	VTVNLQAAE--QGSNVNHLIKVTDQ--SITEGYDDSEGVIAKHAENLIIYDVTFEVDKV 323
Db	389	MNMLSLAAEPNSGKVNKDKIINPTLSLKNNNHANNVWPTSNQOFLNKANYELDDSI 448
Qy	324	KSGDTVTVDIDKNTVPSDLTDSFTI PKIKDNGSEI IATGYVDNKNKOITTYTFTDYVDKYE 383
Db	449	KEGDTFTIKGYQIRPGLLELPAIKQLRSKQSGSVANGVYDKTNTTNTTYTFTNYVDQY 508
Qy	384	NIKAHLKLTSYIDKSKVPNNNTKLDVBYKKTALSSVNTKIIVGYQRPNENRANTANLSMFTN 443
Db	509	NITGSPDLIATPKRETAIKONQNPYMEVTIANEVVKDFIVDYGNKDKNTYT---AAVAN 565
Qy	444	IDPKNHTVEQTIYINPLR-----YSAKETN-----468
Db	566	VDNVNKNHNEWYILNQNNQNPKYAKYFSTVKNGEFIPGEVKVYEVTDTNAMVDSFNPLN 625
Qy	469	-----VNISNGND-----EG 478
Db	626	SSNVKDVTSQAFKVSADGTRVDINFARSWANGKKYIVTQAVRPTGTVYTEYWLTRDG 685
Qy	479	STIIDD-----STII-----KVYKVGD-----NON-----498
Db	686	TTTNDIFYRGTKSTTVTYLNGSSTAQDNFTYSLGDVYWLDDKNKNGVQDDDEKGLAGVYV 745
Qy	499	-LPDSN-----RI-----YDYSVEBVT-----NDYQAQIGNNNVDN-----529
Db	746	TLKDSNNRELQRYVTTQDSGHYQDNLNGTGYTTFVEFALPDNTYFSPANNSTDAIDSDGER 805

Search completed: October 5, 2004, 20:48:21
Job time : 138.638 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:41:22 ; Search time 122.638 Seconds
(without alignments)
2809.455 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKGNLTKKKPIANKSN.....FAGLGALLGKKRKKNKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2 070022	070022 staphylococ
2	5440	96.4	1056	16 Q8CQ72	Q8CQ72 staphylococ
3	4485	79.4	931	2 Q9K113	Q9K113 staphylococ
4	2756	48.8	1171	2 Q9KWX6	Q9KWX6 staphylococ
5	2698.5	47.8	1166	2 Q86489	Q86489 staphylococ
6	2589	45.9	1141	16 Q99W46	Q99W46 staphylococ
7	2589	45.9	1141	16 Q8NXX5	Q8NXX5 staphylococ
8	2586	45.8	1141	16 Q932F7	Q932F7 staphylococ
9	2344.5	41.5	1633	16 Q8CMP4	Q8CMP4 staphylococ
10	2303.5	40.8	1733	2 Q9K114	Q9K114 staphylococ
11	2234	39.6	1385	16 Q99W47	Q99W47 staphylococ
12	2093.5	37.1	953	16 Q99W48	Q99W48 staphylococ
13	2075.5	36.8	955	16 Q8NXX7	Q8NXX7 staphylococ
14	2037	36.1	1347	16 Q8NXX6	Q8NXX6 staphylococ
15	2023.5	35.8	1947	2 Q86487	Q86487 staphylococ
16	2006	35.5	1893	2 Q8KWM1	Q8KWM1 staphylococ

17	1880	33.3	1315	2	O86488	O86488 staphylococ
18	1652	29.3	589	16	Q99VJ4	Q99VJ4 staphylococ
19	1646.5	29.2	946	16	Q8NXX1	Q8NXX1 staphylococ
20	1628	28.8	933	2	Q53653	Q53653 staphylococ
21	1610	28.5	913	2	O86476	O86476 staphylococ
22	1581	28.0	935	16	Q932C5	Q932C5 staphylococ
23	1573	27.9	907	16	Q8NUL0	Q8NUL0 staphylococ
24	1536	27.2	877	16	Q99R07	Q99R07 staphylococ
25	1471	26.1	881	2	Q93MH7	Q93MH7 staphylococ
26	1216	21.5	1698	2	Q91C00	Q91C00 staphylococ
27	1203.5	21.3	3360	16	Q88XB6	Q88XB6 lactobacill
28	785	13.9	882	16	Q92AK9	Q92AK9 listeria in
29	782.5	13.9	903	16	Q8Y697	Q8Y697 listeria mo
30	747.5	13.2	2020	5	Q81D80	Q81D80 plasmodium
31	671.5	11.9	487	2	Q9K112	Q9K112 staphylococ
32	670.5	11.9	970	11	Q8VBY1	Q8VBY1 rattus norv
33	654.5	11.6	540	2	Q8KR21	Q8KR21 staphylococ
34	645.5	11.4	2310	16	Q8CMU7	Q8CMU7 staphylococ
35	640.5	11.3	485	16	Q8CNM7	Q8CNM7 staphylococ
36	639	11.3	579	13	Q7T138	Q7T138 brachydanio
37	624.5	11.1	2271	16	Q99QY4	Q99QY4 staphylococ
38	622	11.0	968	5	Q815W9	Q815W9 plasmodium
39	616.5	10.9	2275	16	Q8NUJ3	Q8NUJ3 staphylococ
40	612	10.8	2283	2	Q8VQ99	Q8VQ99 staphylococ
41	607	10.8	2900	5	Q812E0	Q812E0 plasmodium
42	585	10.0	1182	5	Q81D30	Q81D30 plasmodium
43	560	9.9	970	16	Q8DYL7	Q8DYL7 streptococc
44	554	9.8	3394	5	O77384	O77384 plasmodium
45	549	9.7	2570	2	Q92FF9	Q92FF9 streptococc

ALIGNMENTS

RESULT 1

070022 ID 070022 PRELIMINARY; PRT; 1092 AA.

AC 070022; 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Fibrinogen-binding protein precursor.
 OS Staphylococcus epidermidis.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB;
 RX MEDLINE=98261511; PubMed=9596732;
 RA Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;
 RT "A Fibrinogen-binding protein of Staphylococcus epidermidis.";
 RL Infect. Immun. 66:2666-2673(1998).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
 CC AN AMIDE BOND (BY SIMILARITY).
 DR EMBL; Y17116; CAA76638.1; -.
 DR FIR; T30214; T30214.
 DR GO; GO:0005618; C:cell wall; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR008454; Cna B.
 DR InterPro; IPR005877; GpoS YSIRK.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR Pfam; PF05738; Cna B; 2.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF04650; YSIRK signal; 1.
 DR TIGRfams; TIGR01167; LPXTG anchor; 1.
 DR TIGRfams; TIGR01168; YSIRK signal; 1.
 DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
 DR Cell wall; Peptidoglycan-anchor; Signal.
 FT SIGNAL 1 51
 FT CHAIN 52 1092 FIBRINOGEN-BINDING PROTEIN.
 SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;

Query Match 100.0%; Score 5646; DB 2; Length 1092;

Best Local Similarity 100.0%; Pred. No. 1.7e-185; Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYVK 60
Db	1	MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYVK 60
Qy	61	DSNTDDELSDSDSDEEKNDVNNQSIINTDNNQIIKKEETNNYDGIKRSERDTS 120
Db	61	DSNTDDELSDSDSDEEKNDVNNQSIINTDNNQIIKKEETNNYDGIKRSERDTS 120
Qy	121	TTNVDENEATFLQKTPQDNTLHTEBEVKESSESSVSSNIDTAQOPSHHTINREESVQTS 180
Db	121	TTNVDENEATFLQKTPQDNTLHTEBEVKESSESSVSSNIDTAQOPSHHTINREESVQTS 180
Qy	181	DNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTINDEKISNQDE 240
Db	181	DNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTINDEKISNQDE 240
Qy	241	LLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSTIEGYDDSE 300
Db	241	LLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSTIEGYDDSE 300
Qy	301	GVIKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGEIIA 360
Db	301	GVIKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGEIIA 360
Qy	361	TGYDNNKKNQIYTFDYYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
Db	361	TGYDNNKKNQIYTFDYYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
Qy	1021	SDSGSDSDSDSDSDSLGNSDKSTYKDKLPDGTGANEDYSGKGTLLGTLFAGLGALL 1080
RESULT 2		
Q8CQ72	Q8CQ72	PRELIMINARY; PRT; 1056 AA.
AC	Q8CQ72	(T-EMBLrel. 23, Created)
DT	01-MAR-2003	(T-EMBLrel. 23, last sequence update)
DT	01-OCT-2003	(T-EMBLrel. 25, last annotation update)
DE	Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein.	
DE	Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein.	
GN	SE0331	
OS	Staphylococcus epidermidis.	
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.	
OX	NCBI_TaxID=1282;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 12228;	
RA	Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,	
RA	Chen Z., Wen Y.;	
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB016745; AAC03928.1;	
DR	GO; GO:0016020; C:membrane; IEA.	
DR	InterPro; IPR008454; Cna_B.	
DR	InterPro; IPR005877; Gpos_Ysirk.	
DR	InterPro; IPR001899; Gram_pos_anchor.	
DR	Pfam; PF05738; Cna_B; 2.	
DR	Pfam; PF00746; Gram_pos_anchor; 1.	
DR	Pfam; PF04650; Ysirk_signal; 1.	
DR	TIGRFAMs; TIGR01167; LPXTG_anchor; 1.	
DR	TIGRFAMs; TIGR01168; Ysirk_signal; 1.	
DR	PROSITE; PS0847; GRAM_POS_ANCHORING; 1.	
KW	Complete proteome.	
SQ	SEQUENCE 1056 AA; 115728 MW; 015869A9E5CA2723 CRC64;	
Query Match 96.4%; Score 5440; DB 16; Length 1056;		
Best Local Similarity 96.7%; Pred. No. 1.8e-178;		
Matches 1056; Conservative 0; Mismatches 0; Indels 36; Gaps 1;		
Qy	1	MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYVK 60
Db	1	MINKKNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYVK 60
Qy	61	DSNTDDELSDSDSDEEKNDVNNQSIINTDNNQIIKKEETNNYDGIKRSERDTS 120
Db	61	DSNTDDELSDSDSDEEKNDVNNQSIINTDNNQIIKKEETNNYDGIKRSERDTS 120
Qy	121	TTNVDENEATFLQKTPQDNTLHTEBEVKESSESSVSSNIDTAQOPSHHTINREESVQTS 180
Db	121	TTNVDENEATFLQKTPQDNTLHTEBEVKESSESSVSSNIDTAQOPSHHTINREESVQTS 180
Qy	181	DNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTINDEKISNQDE 240
Db	181	DNVEDSHVSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTINDEKISNQDE 240
Qy	241	LLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSTIEGYDDSE 300
Db	241	LLNLPINEYENKARPLSTTSAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQSTIEGYDDSE 300
Qy	301	GVIKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGEIIA 360
Db	301	GVIKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIKNSGEIIA 360
Qy	361	TGYDNNKKNQIYTFDYYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
Db	361	TGYDNNKKNQIYTFDYYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTALESVVK 420
Qy	421	TITVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDEGST 480

Db	421	TIIVTQRPENRTANLQSMFTNIDTKNHTVEQTIINPLRYSAKETNWNISGDEGST	480
Qy	481	IIDDSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYQAQLGNNDVNFNIDSPYI	540
Db	481	IIDDSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYQAQLGNNDVNFNIDSPYI	540
Qy	541	IKVISKYPNKKDYTTIQOTVTMOTITNEYTGERTASVDNTIAFSTSSGQGGDLPPK	600
Db	541	IKVISKYPNKKDYTTIQOTVTMOTITNEYTGERTASVDNTIAFSTSSGQGGDLPPK	600
Qy	601	TYKIGDVWEDVDKQIGTQNTDNKPLSNVLVTLYPDGTSKSVRTDEGKYQFDGLKNG	660
Db	601	TYKIGDVWEDVDKQIGTQNTDNKPLSNVLVTLYPDGTSKSVRTDEGKYQFDGLKNG	660
Qy	661	LYTKITFETPEGYPTPLKHSNTNPAALDSEGSNVVVTINGQDDMTIDSGFYQTPKYSLGNY	720
Db	661	LYTKITFETPEGYPTPLKHSNTNPAALDSEGSNVVVTINGQDDMTIDSGFYQTPKYSLGNY	720
Qy	721	VWYDTNKGIOGDEKIGISGVKVTLLKDENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDK	780
Db	721	VWYDTNKGIOGDEKIGISGVKVTLLKDENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDK	780
Qy	781	PSGWTQTTTDSGDDQDADGEEVHTITDHDDFSIDNGYDDESDDSDSDSDSDSD	840
Db	781	PSGWTQTTTDSGDDQDADGEEVHTITDHDDFSIDNGYDDESDDSDSDSDSDSDSD	840
Qy	841	SD	900
Db	841	SD	900
Qy	901	SD	960
Db	901	SD	960
Qy	961	SD	1020
Db	961	SD	1020
Qy	1021	SDSGSD	1080
Db	989	-----SD	988
Qy	1081	LGKRRKRRKKN 1092	
Db	1045	LGKRRKRRKKN 1056	

RESULT 3

ID	Q9K113	PRELIMINARY; PRT; 931 AA.
AC	Q9K113	
DT	01-OCT-2000	(TrEMBLrel. 15, Created)
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)
DE	Fibrinogen-binding protein SdrG.	
GN	SDRG.	
OS	Staphylococcus epidermidis.	
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.	
OX	NCBI_TaxID=1282;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=K28;	
RX	MEDLINE=20340957; PubMed=10878118;	
RA	McCreia K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,	
RA	Speziale P., Foster T.J., Hook M.,	
RT	"The serine-aspartate repeat (Sdr) protein family in Staphylococcus	
RT	epidermidis."	
RL	Microbiology 146:1535-1546 (2000).	
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY	
CC	AN AMIDE BOND (BY SIMILARITY).	
DR	ENBL; AF245042; AAF72510.1; -	
DR	GO; GO:0005618; C:cell wall; IEA.	

DR	GO:0016020; C:membrane; IEA.
DR	InterPro; IPR008454; Cna_B.
DR	InterPro; IPR005877; Gpos_YSRK.
DR	InterPro; IPR001899; Gram_pos_anchor.
DR	Pfam; PF05738; Cna_B; 2.
DR	Pfam; PF00746; Gram_pos_anchor; 1.
DR	Pfam; PF04650; YSRK_signal; 1.
DR	TIGRPFAM; TIGR01167; LPXTG_anchor; 1.
DR	TIGRPFAM; TIGR01168; YSRK_signal; 1.
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW	Cell wall; Peptidoglycan-anchor.
SQ	SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 79.4%; Score 4485; DB 2; Length 931;
Best Local Similarity 81.1%; Pred. No. 6.4e-146;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

Qy	4	KKNLLTKKKPANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDQVKDSN	63
Db	3	KKNLLTKKKPANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDQVKDSN	62
Qy	64	TDELSDSNDQSDDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSDETRSTTN	123
Db	63	MDELSDSNDQSSNEEKNDVINNSQSIINTDDNQ-IKKEETNSDAIENRSDITQSTTN	121
Qy	124	VNEATFLQKTPQDNTLHTEBEVKESSVESNSSIDTAQQPSHTTINREESVQTSNV	183
Db	122	VNEATFLQKTPQDNTLHTEBEVKESSVESNSSMDTAQQPSHTTINSEASIQTSNE	181
Qy	184	EDSHVSDPANIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELLN	243
Db	182	ENSRVSDPANIKIESNTESNKEENTIEQPNKVEDSIISQPSYKNIDEKISNQDELLN	241
Qy	244	LPINEYENKARPLSTTSAQPSIKRVTVNQLAAPQGSNVNHLIKVTPQSITEGYDDSEGI	303
Db	242	LPINEYENKVRPLSTTSAQPSISKRVTVNQLAAPQGSNVNHLIKVTPQSITEGYDDSGII	301
Qy	304	KHDAENLIYDVTFFVDDKVKSGDNTVDIDKNTVPSDLTDSFTIPIKIDNSGEIATGT	363
Db	302	KHDAENLIYDVTFFVDDKVKSGDNTVNDKNTVPSDLTDSFPAIKIDNSGEIATGT	361
Qy	364	YDNKNQITVTFDYYVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTLASVKNKIT	423
Db	362	YDNTNKQITVTFDYYVDKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKTLASVKNKIT	421
Qy	424	VEYQRENRTANLQSMFTNIDTKNHTVEQTIINPLRYSAKETNVIISGNGDEGSTIID	483
Db	422	VEYQRENRTANLQSMFTNIDTKNHTVEQTIINPLRYSAKETNVIISGNGDEGSTIID	481
Qy	484	DSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYQAQLGNNDVNFNIDSPYIIV	543
Db	482	DSTIIKVKYKVDNQNLPSNRIDYSEYEDVTNDYQAQLGNNDVNFNIDSPYIIV	541
Qy	544	ISKYDPNKDDYTTIQOTVTMOTITNEYTGERTASVDNTIAFSTSSGQGGDLPPKTYK	603
Db	542	ISKYDPNKDDYTTIQOTVTMOTITNEYTGERTASVDNTIAFSTSSGQGGDLPPKTYK	601
Qy	604	IGDYVWEDVDKQIGTQNTDNKPLSNVLVTLYPDGTSKSVRTDEGKYQFDGLKNGLTY	663
Db	602	IGDYVWEDVDKQIGTQNTDNKPLSNVLVTLYPDGTSKSVRTDEGKYQFDGLKNGLTY	661
Qy	664	KITFETPEGYPTPLKHSNTNPAALDSEGSNVVVTINGQDDMTIDSGFYQTPKYSLGNYVWY	723
Db	662	KITFETPEGYPTPLKHSNTNPAALDSEGSNVVVTINGQDDMTIDSGFYQTPKYSLGNYVWY	721
Qy	724	DTNKGIOGDEKIGISGVKVTLLKDENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDKPSG	783
Db	722	DTNKGIOGDEKIGISGVKVTLLKDENGNIISTTTTDENGKYQFDNLNSGNYIIVHFDKPSG	781
Qy	784	MTQTTTDSGDDEQDADGEEVHTITDHDDFSIDNGYDDESDDSDSDSDSDSDSDSDS	843
Db	782	MTQTTTDSGDDEQDADGEEVHTITDHDDFSIDNGYDDESDDSDSDSDSDSDSDSDS	834


```
RESULT 5
086489
ID AC 086489 PRELIMINARY; PRT; 1166 AA.
AD 086489;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Sdr E protein.
GN SDR E.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=99098700; PubMed=9884231;
RA Joseason E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-apsartate repeat protein multigene
RT family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL: AJ005647; CAA06652.1; -.
DR PIR: T28680.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna B.
DR InterPro: IPR005877; GpoB_YSIK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF05738; Cna B; 3.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YSIK signal; 1.
DR TIGRfam: TIGR01167; LpxTG anchor; 1.
DR TIGRfam: TIGR01168; YSIK signal; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;

Query Match 47.8%; Score 2698.5; DB 2; Length 1166;
Best Local Similarity 48.2%; Pred. No. 8e-85;
Matches 595; Conservative 158; Mismatches 267; Indels 215; Gaps 26;

QY 1 MINKKN-NLTATKKPIANKSNKYAIRKPTVGTASIVIGATLFLGLGHNEAKAEENSVDV 59
DB 1 MINRDNKKAITKGMISNRLNKFIRKYTVGTASIVIGATLFLGLGHNEAKAEENSVDV 58
QY 60 KDSNTDDELSDNQSDDEKNDVINNNQSDINTDDNNQIIKKEETNNYDIEKESDRTE 119
DB 59 ENAKQDATTSDNKEVVSSTENNSTNNP-----IKKG--TNTDSQPEAKKES 110
QY 120 STTNVDNEATFLQKTPDQNTLTHEEVKES---SSVESNSIDTAQPSHTTINREES 176
DB 111 SSTQKQNNVATATETKPN--TEKENVKSTKTATEDTSVILEEKAENNT--NNDVT 166
QY 177 VQTSNDNVEDSHVDFANSKIKESNT---ESGKEENTIEQPNKVBDSSTTSQPSGYTNIDE 233
DB 167 TKPS-----TSEPTSEIQKPTTPQESTNIENSQPPTPSKVD---NQVTDATNPKE 216
QY 234 KIS-NQDELNLNLP-----INYEKARPLSTTSQAQSIKRV-----TWNQAAEQ 277
DB 217 PVNVSEELKNPEKLKELVRNDSNTDHPKVPATAPTSVAPKRVAKRFAVAQPAVA 276
QY 278 GSNVHLIKYTDQSIITGEYDDESEGVIAKHAENLIYDVTFEVDDKVKSGDTMTVDIKNT 337
DB 277 SNNVNDLIKTKYIKVG-DKONVAAAHGKOLEYETETIDNKVKKGDTMTINVDKNV 335
QY 338 VPSDLTDSFTPIKIKNSGELIATGYDNKNKQIYTTFTDYVDKYENIAKHLKLTYSIDK 397
DB 336 IPSDLTDKNPIDITDPSGVIAGTGDATKQIYTTFTDYVDKYEDIKSLRLTYSIDK 395
QY 398 SKVPNNNTKLDVEYKLTALSSVNTKITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYI 457
```

RESULT 6

Q99W46

```
ID AC Q99W46 PRELIMINARY; PRT; 1141 AA.
AD Q99W46;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDR OR SA0521.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
```


[illegible]

Dd	111	STQOQQNNVTAITETKPN--I-EKENVKPSTDKTATEDTSVILEKKAPNT--NNDVT	167
Qy	177	VQTSNDVEDSHVDSFANSKIKESNT---BSGKEENTIEQPNKVKEDSTTSQPSGYTNIDE	233
Dd	167	TKPS-----TSEIQTKPTTPOBSTNIENSQOPTSKVD---NQVTDATNPKE	211
Qy	234	KIS-NQDELLNLP-----INENYKARPLST--TSAQP-----SIKRVTVNQLAAEQ	277
Dd	212	PVNSKBEELKNPEKLKELVRNDNNTDRSTKEVATAPTSVAPKRLNAKRFAPAQAAVA	271
Qy	278	GSNVNHLIKVTDQSITEGVDSEGVKHAHAENLJYDVTFEVDVKVSGDMTMTVDIDKNT	337
Dd	272	SNVNDLITVTKIKVG-DGKNVAAAHGKDIEYDTEFTIDNKKGDMTMIINDKNV	330
Qy	338	VPSDLTDSFTPKIKDNNGSEIIATQYDNKNKQIITYTFTDYVDKYENIKAHLKLTYSIDK	397
Dd	331	IPSDLTDKNRPIIDTDPGSEVIAKGTFOKATKQITVTFTDYVDKYEDIKARLTLSYIDK	390
Qy	398	SKVPNNNTKLDVEYKTAJLSSVYNKTLITVEYORENERTANLOSMFTNIDTKHTEVQTIYI	457
Dd	391	QAVP-NETSLMLTFAATKETSQNVSVDYQDPWVHGDSNIOISIFTKLENKQTIIEQQIYV	449
Qy	458	NPLRYSAKETNWNISGNG-----DEGSTIIDDSTIIKVKYKGDQNLPDSNRIYDY	508
Dd	450	NPLKKTATNTKVDIAGSQVDYDGNIKLNGSTIIDQNTIEIKVKYKVPNQLPQSNRIYDP	509
Qy	509	SEYEDVTND-DYALQGNNDVNIAPGNTDSPIYIKVISKYDPNKDDYTTIOQVTMOTTI	567
Dd	510	SQYEDVTSQFNKKFSFNNAVTLDFGINSAYIIKWSKYTPTSDEGLDIAQGTSMRITT-	568
Qy	568	NEYTGEFRASDYNTIAFSTSQQOQDGL-PPEKTYKIGDYVWEDVDKXGQNTNDNKEP	626
Dd	569	DKY-GYNYAGYSNFIVTSNDTGGDGVTKPEEKLYKIGDYVWEDVDKGVQGTDSKEKP	627
Qy	627	LSNVLVLTYPDGTSKSVRTDEBCKYQPDGLKNGLTYYKITPTEPEGYPTPLKHSGTNPAL	686
Dd	628	MANVLVLTYPDGTYSVRTDANGHYEFGGLKOGETTVKFTPAGLPTKVNGTGTDGEX	687
Qy	687	DSEGNYSVMVTINGOOD-----	702
Dd	688	DSNGSSITVKINGKDDMSLDTGFYKPEKYNLGDYVWEDTNKDGIQDANEPGIDKVVTLK	747
Qy	703	-----	702
Dd	748	DSTGKVI GTTTTDAAGKYKFTDLDNNGYNTVEFETPAGYTPVTKNTAEDKXDSNGLTTTGV	807
Qy	703	-----MTIDSGFYCTPKYSLGNVYVYDNTKDGIOGDEKGISGVKVLTKDENGLI STT	756
Dd	808	IKDADNWTLSGFFKTPKYSGLCDYVWDSNKGQDQDSTERGIKOVKVTLLNKEGEVIGTT	867
Qy	757	TTDBNGKYQFDNLNSGNYI VHFDPKSGVOTTTDSGDDEQADAGEEYHVHTITDHDFFSI	816
Dd	868	KTDENGKYRFDNLDSGKYKVI FEKAPAGLTQTVNTTDEDD-KDADGGEVDVITDHDHDFIL	926
Qy	817	DNGYVYDDESDDSD	876
Dd	927	DNGYFEEDT-----SD	955
Qy	877	SD	936
Dd	956	SD	1015
Qy	937	SD	996
Dd	1016	SD	1075
Qy	997	SD	1055
Dd	1076	SDSDSDAGKHTPVKPM-----TTKDHNNKAKALPE	1106
Qy	1056	TGANEDYGSXGTLTGTLFAGLGAALLL-GKRRKNRK	1089
Dd	1107	TGSENNGSNNATLFGGLFAALGSLLLFGRKKQNK	1141

```
Db 472 NITGSEDLIATPKRETAIKDNQNYPMEVTIANEVVKDFIVDYGKKNKNTTT---AAVAN 528
Qy 444 IDTKNHTVEQTIYINPLR-----YSAKETH-----468
Db 529 VDNVANKNEVYLNQNNQPKYAKYFSTVKGKPIPGEVKVYEVTDITNWMVDSFNPDLN 588
Qy 469 -----STII-----KVKVGD-----VNISGND-----EG 478
Db 589 SSVKDVTSQTPKVSADGTRVDINPARSMANGKKYIVTQAVRPTGTGNVYTYEYWLTRDG 648
Qy 479 STIIDP-----STII-----KVKVGD-----NON-----498
Db 649 TTNWDFYRGTSTVITYLNGSTAQDNPTVSLGDIYVWLDKRVKGVQDDDEKGLAGVYV 708
Qy 499 -LPDSN-----RI-----YDYSEYEDVT-----NDYIAQLGNNDVN-----529
Db 709 TLKDSNNRELQRTVTDQSGHYQFDNLQNGTYTVEFAIPDNYTPSPANNSTNDAIDSGER 768
Qy 530 -----INFGNIDSPYI-----IKVISKY-----DPNKD-----DYTTIQ 558
Db 769 DGRKVVVAAGTIN--NADNMVTVDGTGYLTPKYNVGVYVEDTNKDGIQDDNEKGISNVK 826
Qy 559 QTV-----TMOITINEYTGERTASYDN---TIAFST-----SSGQGG-DLPE 599
Db 827 VTLKNKNGTIGTTTTDSNGKYEFTGLENGDYTIEFETPEGYTPYKQNGSGDEGKDSNGT 886
Qy 600 KT-----YKIGDYVEDVDKGIQNTNDNEKPLSNVLTLTPYDG 639
Db 887 KTTVTVKADNKTIDSGFYKPIYNLGDYVWEDTNKDGIQ--DSEKGISGVKVTLDKNG 944
Qy 640 TS-KSVRTEDGKYQPDGLKNGLYKLTETPEGYTPTLKHSNTNPDALSGNSVWVTIN 598
Db 945 NAIGTTTTDASGHYQFKGLENG--SYTVRFETPSGYTPYKANSQDITVDSNGITTTGIIN 1003
Qy 699 GDDMTIDSGFYQTPKYSLGNYVYDNTKDGIGDDEKGISGVKVTLDKNGVILSTTTT 758
Db 1004 GADNLITIDSGFYQTPKYSVGVYVWEDTNKDGIQDDNEKGISGVKVTLDKNGVILSTTTT 1063
Qy 759 DENGKYQFDNLNSGNYIVHFDKPSGMTQTTTSGDDDEQADGEEVHVHTITDHDPSIDN 818
Db 1064 DENGKYQFDNLDSGNYIIHFKEPEGMTQTANSNGNDEKADGEDVRVITDHDPSIDN 1123
Qy 819 GYDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 878
Db 1124 GYFDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1183
Qy 879 -----878
Db 1184 SDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDAD 1243
Qy 879 -----878
Db 1244 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDAD 1303
Qy 879 -----878
Db 1304 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDADSD 1363
Qy 879 -----SDSDSDSDSD 888
Db 1364 AUSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSAD 1423
Qy 889 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS-----942
Db 1424 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSADSD 1483
Qy 943 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSSD 1002
Db 1484 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSSD 1543
Qy 1003 SDSDSDSDSVSDSDSDSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1062
Db 1544 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSKNAK 1603
```

```
Qy 1063 GSKGTLGLTFLAGLALLGKRRKRNK 1091
Db 1604 DSKGTLGLTFLAGLALLGRRRK-KDNK 1631

RESULT 10
Q9KI14 PRELIMINARY; PRT; 1733 AA.
AC Q9KI14;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative cell-surface adhesin SdrF.
GN SDRF.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=9491;
RX MEDLINE=20340957; PubMed=10878118;
RA McCreia K.W., Hartford O., Davis S., Ni Bidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis.";
RL Microbiology 146:1535-1546(2000).
DR EMBL; AF245041; AAF72509.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna B.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna B; 4.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
SQ SEQUENCE 1733 AA; 184720 MW; D8D62EA1692FD4E8 CRC64;
```

Query Match 40.8%; Score 2303.5; DB 2; Length 1733;
Best Local Similarity 34.6%; Pred. No. 3.6e-71;
Matches 605; Conservative 146; Mismatches 313; Indels 685; Gaps 46;

```
Qy 2 INKKNLLTKKPIANKSNKYAIRFTVGTASIVGATLLFGLHNEAKAEN-----S 55
Db 9 INKRVDFL-----SNKYNKYSIRKFTVGTASILVGATLMFGAADNEAKAEDNQLESAS 62
Qy 56 VQDVKDS--NTDDELS--DSNDQSDDEKNDVINN-----NQSN 91
Db 63 KEEQKGRDNENSKLNQVDLNGSHSSEKTTWNNAATEVKVEAPTSDVSRPKANEAV 122
Qy 92 TDD-----NNQII---KKEETNNYDGIK-----RSEDRTESTTNVDEN 127
Db 123 TWESTKPKTEAPTVEESIAETPKTSTTQODSTKKNPDLNLSSTTSKESKDEH 182
Qy 128 EATFIQ-KTPQNTHLTBEVK-ESSSVSSNSSIDTAQOPS-----167
Db 183 STKQAMSTNKNLDTNDSPTQSEKTSQANDSDTNQSPAKQLDSKSEQKVKYTKFN 242
Qy 168 -----HTTN-RBESVQTSNDVSDSHVDFANSKIKESNTSEKENTIEOPNKVKE 218
Db 243 DEPTQDVEHTTKTPFSVSTDSVNDK--QDYTSAV-----ASLGVDSENEITAINVR 296
Qy 219 DSTTSQPSGYTNIDRKI-----SNODELNLNLP--INEYENKARPLSTTSAQPSIKR 267
Db 297 DNLDLKAASREQINEAIIAEALKKDFNPDYGVDTPLALNRQSKNSP--HKSASE---R 351
Qy 268 VTNQLAAB--QGSNVNHLIKVTDQ--SITEGYDSEGVKAHDAENLYDVTFFVDDKV 323
Db 352 MNLMSLAEPNSGKNVNDKVKITNPTLSLNKNHANNVWPTNSNQFNLKANYELDDSI 411
Qy 324 KSGDWTWTDIDKNTVPSDLTDSFTPIKIKDNGSEIITATGYDNKNQKITVTTFTDYDYKVE 383
Db 412 KEGDTFTIKYQYIRPGLLEPAIKQLRSKGSIVANGVYDKYTKTNTTFTTYNYDQYQ 471
```



```
Qy 120 STNVNEN-EATFLOKTPQDNTHLTBEEVKESSESSVSSNIDTAQPSHTTINREESVQ 178
Db 103 TTSNGNKSIEKESVQSTTGKNEVSTAKSDEQAPSTNEDLTKQ-----TISNOEGLQ 157
Qy 179 TSNVEDSHVSDPANSKIKESNTESGKEENTIQPN-KVKEDSTTQPSGYTIDSKISN 237
Db 158 -PDLENKSVN-----VQPTNEENKVKDAKTESITLNVKSDAISK-----NAETLVDN 205
Qy 238 QDELLNLPINEYENKARPLSTTGAQPSIKRVTVNQLAAEQGS-----NVNHLIKVTDQGIT 293
Db 206 NSNSNN-----ENNADIILPKSTAPKSLNTRMRMAAIQPNSTDSKNVNDLITSNTTLTV 259
Qy 294 EGYDDSEGLVKADAENLIYDVTFEVDVKVSGDTWTDIDKTV-----PSDLTDSFT 347
Db 260 VDADNSKTIIVPAQDYLKSLKQIT--VDDKVKSGDYFTIKY-SDTQVQYGLNPEDIKN--- 313
Qy 348 IPKIKD-NSGEIATGYDNKNKQITVTFDYDDKYENIKAKHLKLSYIDKSKVPNNNTK 406
Db 314 IGDIKDPNGETIATAKHTANLLITYTFDYDRNSVQMGINSYIYMDADTIPVD--K 371
Qy 407 LDVEYKALSSVNKTIITVEYQRP--NENRTANI-QSMPTNIDTKNHT-----VSEQTIY 456
Db 372 KDVPFSVTIGNQITTTTADITYPAYKEADNNSIGSAFT--ETVSHVGNVEDPCGYNQVY 429
Qy 457 INPLYSAKETNWNISNGDEGST-----IIDDSTIIKVKYKVGDNQMLPDSNRIDY--SE 510
Db 430 VNPMDKDLGAKLKVAYHPKYPTNIGQINQNVNTNIIKVRPPEGYTL---NKGYDVTND 486
Qy 511 YEDVTND--DYAQLGNNDVNINFGNIDSPYIIKVIKSKYDPNKDDVTYIQQVTMTOTTIN 568
Db 487 LVDVTDEFKRWYTGNSQSNVLDGDLTSAYVWVNTKFOYTNSPFTLVQMATLSTGN 546
Qy 569 EYTGFEPTASYDNTIAPSTSSGQGGDLPPKTYKIGDYWEDVDKDGIO----- 618
Db 547 -----KSVSTGNALGFTNNQSGAG---QEVYKIGNVYEDTNKNGVQELGEKGVGVN 596
Qy 619 -----NTN----- 621
Db 597 TVTVFDNNTNTKVGAEVTKEDGYLIPNLPNGDYRVEFNSLPKGYEVTFPSKQGNBELDS 656
Qy 622 -----DN-----EKPLSNVLVTI----- 634
Db 657 NGLSSVITVNGKNLSADLGIYKPKNLGYWEDTNKNGIQDQDEKSGISGVTVILKDN 716
Qy 635 -----TYPDG----- 639
Db 717 GNVLKTVTTDADGKYKFTDLNGYKVEFTTPGTYPTTTSKDIKDSNGLTITGVIN 776
Qy 640 -----TSKSV-----RT 646
Db 777 GADNMTLDSGFYKTPKYNLGNVYWEDTNKDGKQDSTEKGISGVTVTLKNGENGEVLQTTKT 836
Qy 647 DEBGKYQDGLKNGLTAKITFETPEGVYPTLKHSGTNPALDSEGS----- 692
Db 837 DKDGKYQFTGLENG-TYKVEFETPSGYTPTQVSGTDEGIDSGTSTGVIKDKNDTID 895
Qy 693 -----VTI-----NGQ----- 700
Db 896 SGFYKPTYNLGDYWEDTNKNGVQDKDEKGISGVTVTLKDNKVLKTVTTDENGKYQFT 955
Qy 701 -----DMTIDSGFYQTPKYS 716
Db 956 DLNNGTYKVEFETPSGYTPTSVTSGNTEKDSNGLTITGVIKADNNTLDSGFYKTPKYS 1015
Qy 717 LGNVYVWDTNKGIOGDDEKIGISGVKTVLKDENGNISITTTDENGKYQPDNLNSGYIV 776
Db 1016 LSGYVWYDSNKGQDQSTEGIKDKVRKILLNEXGEVIGTITDENGKYRFDNLDSGKYK 1075
Qy 777 HFKPSGMTQTTTDSGDDQDADGEVHYTITDHDDFSIDNGYDDE--SDSDSDSDSDS 835
Db 1076 IFKPTGLTGTGNTTDED--KDADGGEVDVTIITDHDFTLUNGYYEEETSDSDSDSDS 1134
Qy 836 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 895
```

```
Db 1135 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1194
Qy 896 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 955
Db 1195 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1254
Qy 956 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1015
Db 1255 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1309
Qy 1016 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1064
Db 1310 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1359
Qy 1065 KGTLLGTLTLAGLALL-GKRRNRNK 1089
Db 1360 NATLFGFLFAALGSLLLFGRKKQNK 1385
RESULT 12
Q99W48 PRELIMINARY; PRT; 953 AA.
AC Q99W48,
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDRC OR SAV0561 OR SA0519
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Staphylococcus aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kato M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayaishi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF003359; BAB56723.1; -.
DR EMBL; AF003131; BAB41750.1; -.
DR PIR; C89824; C89824.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna B; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfam; TIGR01167; LPXTG anchor; 1.
DR TIGRfam; TIGR01168; Ysirk signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;
Query Match 37.1%; Score 2093.5; DB 16; Length 953;
Best Local Similarity 43.8%; Pred. No. 3e-64;
Matches 488; Conservative 140; Mismatches 298; Indels 189; Gaps 28;
Qy 2 INKNNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGHNEAKAEENSVDQVKD 61
Db 1 MNKTKATNRKGMIPNRLNFKSIRKYSVGTASIVLTGTLFGLSGHEAKAAE----- 52
Qy 62 SNTDDELSDNSDQSDSEKNDVINNNQSNINTDNNQIIKKETNNYDIEKRSEDTTEST 121
```

```
Db 53 -HTNGELNQSNETTAPSEN-----KTT 74
Qy 122 TNVDENEATFLOKTPQDNTHLTREEVKSESSVSSNSSIDTAQOPSHHTTINREESVQTS 181
Db 75 EKVDSRQL-----KONTQTA-----TAQPKVT-----97
Qy 182 NVEDSHVDFPANGKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEX---ISNQ 238
Db 98 -----MSDSATVKETSSNQSS-----PQNATASQSTTQTSNVTNDKSTTYSNE 142
Qy 239 DELLNLPINEYEN-KARPLSTTSAQPSIKRVTVNQALAA-EQGSNVNHLIKVTQSI-----292
Db 143 TDKSNL--TQAKVSVTPKTTIKPRALNRMVNTVAAPQOGTNNVDKVFHTNIDIAIDK 201
Qy 293 -----TEGVDDSGVIAKHAENLIYDVTFEVDKVKSGDVTMTVDIDKNTVPSDLT 343
Db 201 GHVNTKNTGFATSSDVLK-----LKANYTTIDDSVKEGDTFTFKYQGFYRGSVR 252
Qy 344 DSFTIPKIDNSGEIATGYDNKNKQITTYTDDYVDKYENIKAHKLKLTYSIDKSKVPNN 403
Db 253 LPSQTNLYNAQGNIIAKGIYDSKNTTNTTYTNYVDQYTNVSGSPQVAFAKRENATTD 312
Qy 404 NTKLDVEYKTALSUNKTIIVFYQRPENRNTANLQSMFTNIDTNKHTVEQTIYN-PLRY 462
Db 313 KTAYPEVTLGNDTYSKDVIVDY---GNQKGQOLISSTNYINNEDLSRNMNTVYVNOPKKT 369
Qy 463 SAKETNV-NISGNGDEGSIIDSTIIKVKYVGNQNLPSNRIDYSEVEDVTND-DYA 520
Db 370 YTKETFTVNTU-----GYKNPDAKPKIYEVDQNFQVDS-FTPDTSKLDKVGQFDVI 423
Qy 521 QLGNNDVNIINFGN-----IDSPYIIKVISYDPNKDDYTTIQOTVTMTQTTINEYTGFEFT 576
Db 424 YSNDKNTATVDLLNGQSSSKQYIIQQVAYPDNSSTDNGKIDYTLQNGKSSWSN---479
Qy 577 ASDNTIATFSSGQGGDLPPEKTYKIGDYVEDVDKQIQTNDNKRPLSNVLTLY 636
Db 480 -SYSNVNGSSTANGD-----QKYNLGDYVWEDTNKQKQDA--NEKGIKGVYVILKD 529
Qy 637 PDGTS-KSVRTDGDQYQDGLKXGLTKITFTPEGYTPTLKHSGTNPALDSEGNSSVMV 695
Db 530 SNGKELDRITTDENGKYQFTGLNG-TYSVFSTPAGYFTTANAGTDDAIVSDGLHTTG 588
Qy 696 TINGQDDMTIDSGYOPPKYSLGNYVMYDNTKQGIQGDDEKGISGVKVTILKDENGNIIST 755
Db 589 VIKADNMTLDSGFYKTPKYSLGDYVMYDSNKGDKQDSTEGIKGVKVTILQNEKGEVIGT 648
Qy 756 TTTDENGKYQFDNLNSGNIYVHFDKPSGWTQTTTDSGDDDEQADGEEVHTTIDHDDPS 815
Db 649 TETDENGKYRFDNLDSGKYKVIPEKPAGLTQGTNTTTEDD-KDADGGEVDVTTIDHDDFT 707
Qy 816 IDNGYDDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db 708 LDNGYEEET-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 766
Qy 876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db 767 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 826
Qy 936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db 827 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 886
Qy 996 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1055
Db 887 DSDSDSAGKHTTKPMSTV-----KQOHTAKALPE 918
Qy 1056 TGANEDYGSKGLTGLTFLAGLALL-GKRKXNRK 1089
Db 919 TGSNNNSNNGTLFGGLFAALGSLLLFGRKXQNK 953
```

RESULT 13

```
Q8NXX7
ID Q8NXX7 PRELIMINARY; PRT; 955 AA.
AC Q8NXX7;
DT 01-OCT-2002 (Tr-EMBLrel. 22, Created)
DT 01-OCT-2002 (Tr-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (Tr-EMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding bone sialoprotein-binding
DE protein.
GN SDR OR MW0516.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi P., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004823; BAB94381.1; -.
DR GO; GO:0016020; C-membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF05738; Cna_B; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXFG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ
SEQUENCE 955 AA; 103642 MW; 65FD8BB93A477A22 CRC64;

Query Match 36.8%; Score 2075.5; DB 16; Length 955;
Best Local Similarity 43.8%; Pred. No. 1.2e-63;
Matches 485; Conservative 144; Mismatches 307; Indels 171; Gaps 27;

Qy 2 INKONLLKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGHNEAKAEENSVDVKD 61
Db 1 MNKKTVTRKGMIPNRLNKFIRKYSVGTASILVGLTILFGLSGHEAKAE-----52
Qy 62 SNTDDELSDNSQSSDEEKDYNNNNQSNINTDDNNQI IKKEETNNYDIEKSEDETEST 121
Db 53 -HTNGELNQSNETTAPSEN-----KTT 74
Qy 122 TNVDENEATFLOKTPQDNTHLTREEVKSESSVSSNSSIDTAQOPSHHTTINREESVQTS 181
Db 75 EKVDSHQL-----KONTQTA-----TAQPKVT-----97
Qy 182 NVEDSHVDFPANGKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEX---ISNQ 238
Db 98 -----MSDSATVKETSSNQSS-----PQNATASQSTTQTSNVTNDKSTTYSNE 142
Qy 239 DELLNLPINEYENKARPLSTTSAQPSIKRVTVNQALAA-EQGSNVNHLIKVT--DQSITEG 295
Db 143 TDKSNL-TQAKDVSAATPKTTIKPRALNRMVNTVAAPQOGTNNVDKVFHTNIDIAIDK 201
Qy 296 Y---DDSEGVIAKHAENLIYDVTFEVDKVKSGDVTMTVDIDKNTVPSDLTDSFTPIK 352
Db 202 HLNKDTGKTEFWATSSDVLKANYTTIDDSVKEGDTFTFKYQGFYRGSVRPSQTNLY 261
Qy 353 DNSGEIATGYDNKNKQITTYTDDYVDKYENIKAHKLKLTYSIDKSKVPNNNTKLDVEYK 412
Db 262 NAGENLIAGIYDSTNTTNTTYTNYVDQYTNVSGSEFQVAFAKRENATTDKTAKEVS 321
Qy 413 TALSSVNKTIIVFYQRPENRNTANLQSMFTNIDTNKHTVEQTIYN-PLRYSAKETNV-N 470
Db 322 LGNDTYSBEIIVDY---GNKKAQPLISSTNYINNEDLSRNMNTAYVNOPKVTYTKQTFVN 378
Qy 471 ISGNGDEGSIIDSTIIKVKYVGNQNLPSNRIDYSEVEDVTNDYDQALGNNN---D 527
```



```
Db 379 LT-----GYKFNPAKFKIYVTDQNFQVDS--FTPDTSKLKDVTNQFNITYSNDNKAT 432
Qy 528 VNINFGNIS--PYLIKVKISKDFNKDYYTIIQQTVMQTTINNEYTGBFRFASYNWTF 585
Db 433 VDLMGQTSNKKQYIIQQVAYPDNTSTDKGIDYTLDTDKTKYSWEN-----SYSNVNGS 487
Qy 586 STSSGOGQGLDPEKTKYKIGDYVWEDVDKDGQNTNDNEKPLSNVLATLITYPDGTS--KSV 644
Db 488 STANGD-----QKYNIGDYVWEDTNKDGQDA--NEKGIGVYVILKDSNGKELDR 538
Qy 645 RTDEDKYQFDGLKGLTKYITPETPEGYTPTLKHSGTNPALDSEGNVWVWVINGODDMT 704
Db 539 TTDENKQYQFTGLSNG--TYSVEFSTAGYPTTANAGTDDAVDSGLTGTGVKOADNMT 597
Qy 705 ISGFTQTPKISLGNVWVDNKGIOGDDKEKISGVKVTLDKDNENIISTTTTDENGKY 764
Db 598 LDSGFYKTPKYSIGDYVWYDSNKGQDSTKGIKGVKVTLQNEKGEVIGTTTDENGKY 657
Qy 765 QFDNLNGYVHFDPKSPGWTQTTDSGDDDEADGEEVHTITDHDDFSIDNGYDD 824
Db 658 RFDNLDSGKYKVFEPAGLTQGTNTEDD--KADGGEVDVTTDHDFTLNGYEE 716
Qy 825 -SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 883
Db 717 TSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 776
Qy 884 DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 943
Db 777 DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 836
Qy 944 DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1003
Db 837 DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 896
Qy 1004 DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1063
Db 897 GKTPAKPMSV-----KQHKTKAKLPETGSENNNS 928
Qy 1064 SKGTLLGTLPAGLGALL--GKRKNRK 1089
Db 929 NNGTLFGGLFAALGSLLLFGRRKKQNK 955

RESULT 14
Q8NXX6 PRELIMINARY; PRT; 1347 AA.
AC Q8NXX6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Aap rich fibrinogen-binding bone sialoprotein-binding
DE protein.
GN SDRD OR MW0517.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
DR EMBL: AP004823; BAB94382.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR001899; Gram_pos_anchor.
DR pfam: PF05738; Cna_B; 5.
DR pfam: PF00746; Gram_pos_anchor; 1.
DR pfam: PF04650; YSIRK_signal; 1.
```

```
DR TIGRFAMs: TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1347 AA; 145958 MW; AB8D393884AD8034 CRC64;

Query Match 36.1%; Score 2037; DB 16; Length 1347;
Best Local Similarity 36.3%; Pred. No. 3.6e-62;
Matches 529; Conservative 165; Mismatches 286; Indels 476; Gaps 45;

Qy 1 MINKKN-NLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVDV 59
Db 1 MLARENKTAIRKGMVSNRLNKFIRKYTVGTASILVGTTLFGLGNQSAKAESTNKL 60
Qy 60 KDSNTDDELSDNSDSSDEEKVDVNNQSNINTDNNQIIKKEETNNYDGIKRSERDTE 119
Db 61 NEATT-----SASDQSSD--KVDMLQNLQEDNTKNDQ-----KEMVSSQGN 102
Qy 120 STTNVDEN-EATFLQKTPQDNTHLTTEEVEKSSSVESNSSIDTAQPSHTTINREESVQ 178
Db 103 TTSNGKSKTEKESVQSTTGNKVEVSTAKSDEQSPKSTNEDLNTKQ-----TISQEA 157
Qy 179 TSDNYEDSHVSDFANSKIKESNTESGKEBNTIEQPNKVKEDSTTSQPSGYTNIKESIQ 238
Db 158 -PDLQENKSVVNAQPTNEENKVKDAKTESTTL---NVKSDAIKS-----NAETLVD 206
Qy 239 DELLMLPINEYENKARPLSTTSAQP-----SIRKVTNNQLAAEQGSNNVHLIKVTD 294
Db 207 SNSNN-----ENNADIILPKSTAPKRLNTRIRIAAVOPSTEAKNNDLITSTNTLV 260
Qy 295 GYDDSEGVKKAHAENLIYDVTFEVDDDKVSGDVTVDIDKNTV-----PSDLTSFT 348
Db 261 DADKNKIVPAQDYLELSQI--KVDDDKVSGDYFTIKY-SDTVQVYGLNLPEDIKN---I 314
Qy 349 PKIKD-NSGEIITATGYDNKNKQITTYTFTDYVDKYENIKAHILKLSYIDSKVPNNNTKL 407
Db 315 GDKDPNGETATAKHDTANNLIITYTFTDYVDRFNSVQMGINYSIYMDADTIP--VSKN 372
Qy 408 DVEYKALSSVNNKTITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYI 457
Db 373 DVEFNVTIGNTTKTNTANIQYDPYVNEKNSIGSAFT--ETVSHVGNKENPGYKQTIYI 430
Qy 458 NPLRYSAKETVNI-----SGNDEGSTIIDDSTIIKVKYGDNQNLPSDNLRYDYS--EY 511
Db 431 NPSENSLNAKLKQVAYHSSYPNNIQQINKEVTDIKIYQPKGYTL---NKGVDVNTKEL 487
Qy 512 EDVTNDVYQAQ---LGNNDNVINFGNIDSPYIIKVISKDPNKKDDYTTIIQQTVMQTTIN 568
Db 488 TDVTN-QYLQKITYGDNNSAVIDFGNADSAVYVWVNTKFQYTTSESPTLVQNVTLs--- 542
Qy 569 EYTGFRFASYNWTFIAFSTSSGQGDLPPEKTKYKIGDYVWEDVDKDGIO----- 618
Db 543 --SDNSKSASMGNALGFTNNQSGAG---QSVYKIGYVWEDTNKNGVQELGEKGVGNV 596
Qy 619 -----NTN----- 621
Db 597 TVTVFDNNTNKVGEAVTKEDSGYLIPNLPGDYRVFSPNLKPGYEVTPSKQGNNEELDS 656
Qy 622 -----DN-----EKPLSNVLVTI--- 634
Db 657 NGLSSVITVNGKDNLSADLGIYKPKYNLGDYVWEDTNKNGIQDQDEKGISGVTVLKDN 716
Qy 635 -----TPDG----- 639
Db 717 GNVLKVTITDADGKYKFTDLNGYKVEFTTPEGVPTTTSVSGDIEKDSNGLTITGVIN 776
Qy 640 -----TSKSV-----RT 646
Db 777 GADNMTLDSGFYKTPKYNLGNVWEDTNKDGKQDSTEKGISGVTVTLKNENGEVLQTTKT 836
Qy 647 DEDGKYQFDGLKGLTKYITPETPEGYTPTLKHSGTNPALDSEGNs----- 692
Db 837 DKDGKYQFTGLENG--TYKVEFETPSGYTPTQVSGTDEGIDSNGTSTTGVIKDKNDTID 895
```


Db 826 SDS 885
Qy 1001 SDSDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060
Db 886 SDAGKHTPAKPMSTV-----KQHKAKALPETGSEN 917
Qy 1061 DYGSKGTLGTLFAGIGALL-LGKREKNRK 1089
Db 918 NNSNNGTLFGGLFALGSLLSFGRRKKQNK 947

Search completed: October 5, 2004, 20:52:22
Job time : 143.638 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:34:15 ; Search time 22.8315 Seconds
(without alignments)
2490.441 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKKNLLTKKKPIANKN.....FAGLGALLKKRKNRKN 1092

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1319.5	23.4	1637	1 MRSP_STAAU	P80544 staphylococ
2	708.5	12.5	1253	1 DPPP_HUMAN	Q9nzv4 homo sapien
3	675.5	12.0	934	1 DPPP_MOUSE	P97399 mus musculu
4	606.5	10.7	687	1 DPPP_RAT	Q62598 rattus norv
5	554	9.8	406	1 SR40_YEAST	P32583 saccharomyc
6	487.5	8.6	1018	1 FNBA_STAAU	P14738 staphylococ
7	471	8.3	640	1 SAMP_PLAFAW	PI3821 plasmodium
8	388.5	6.9	695	1 ARP_EUGER	Q04732 euglena gra
9	368	6.5	1664	1 SLPI_CLOTH	Q06852 clostridium
10	361.5	6.4	400	1 RTOA_DICDI	P54681 dictyostell
11	343.5	6.1	1658	1 YM67_YEAST	Q03661 saccharomyc
12	335.5	5.9	4910	1 MDN1_YEAST	Q12019 saccharomyc
13	333	5.9	797	1 VG48_HSVSA	Q11033 herpesvirus
14	331.5	5.9	1419	1 ALAI_CANAL	O13368 candida alb
15	315	5.6	937	1 HYRI_CANAL	P46591 candida alb
16	314	5.6	2492	1 ATRX_PANTR	Q7ygm4 pan troglod
17	311	5.5	2452	1 RPB1_PLAFAW	P14248 plasmodium
18	310	5.5	1189	1 YJH6_YEAST	P47035 saccharomyc
19	303	5.4	2492	1 ATRX_HUMAN	P46100 homo sapien
20	301.5	5.3	503	1 DMP1_MOUSE	O55188 mus musculu
21	297	5.3	2492	1 ATRX_PONPY	Q7ygm3 pongo pygma
22	295.5	5.2	489	1 DMP1_RAT	P98193 rattus norv
23	294.5	5.2	429	1 DM41_YEAST	P18899 saccharomyc
24	291	5.2	1744	1 TANA_XENLA	Q01550 xenopus lae
25	288	5.1	389	1 SERI_BOMBO	P07856 bombyx mori
26	285.5	5.1	513	1 DMP1_HUMAN	Q13316 homo sapien
27	283.5	5.0	1337	1 DEXT_STRDO	P39693 screpococ
28	283	5.0	914	1 PBPA_BACSU	P39793 bacillus su
29	276.5	4.9	1085	1 IFH1_YEAST	P39520 saccharomyc
30	276	4.9	500	1 GAR2_SCHPO	P41891 schizosacch
31	271.5	4.8	1025	1 MK21_YEAST	Q12176 saccharomyc
32	271	4.8	688	1 LIP_STAEP	Q02510 staphylococ
33	271	4.8	2476	1 ATRX_MOUSE	Q61687 mus musculu

34	269.5	4.8	3178	1 YS89_CAEEL	Q09624 caenorhabdi
35	267.5	4.7	1070	1 PVDR_PLAVB	P22230 plasmodium
36	267	4.7	2867	1 RBP2_PLAVB	Q00799 plasmodium
37	266.5	4.7	1912	1 VITI_CHICK	P87498 gallus gall
38	265	4.7	593	1 SANT_PLAFA7	Q03400 plasmodium
39	264.5	4.7	667	1 CYLI_BOVIN	P35662 bos taurus
40	264.5	4.7	936	1 PHLI_YEAST	P39521 saccharomyc
41	264	4.7	918	1 YMJB_CAEEL	P34487 caenorhabdi
42	263.5	4.7	1770	1 PMPC_CHLTR	O84419 chlamydia t
43	262.5	4.6	279	1 SPBP_RAT	P87923 rattus norv
44	262.5	4.6	817	1 YG4A_YEAST	P46949 saccharomyc
45	262	4.6	510	1 DMP1_BOVIN	Q95120 bos taurus

ALIGNMENTS

RESULT 1

ID	MRSP_STAAU	STANDARD;	PRT;	1637 AA.
AC	P80544; Q9ZF62;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Methicillin-resistant surface protein precursor.			
GN	PUS.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Isolate 1061;			
RX	MEDLINE=21189215; PubMed=11292719;			
RA	Savolainen K., Paulin L., Westerlund-Wikstrom B., Foster T.J.,			
RA	Korhonen T.K., Kuusela P.;			
RT	"Expression of pls, a gene closely associated with the mecA gene of			
RT	methicillin-resistant Staphylococcus aureus, prevents bacterial			
RT	adhesion in vitro.";			
RL	Infect. Immun. 69:3013-3020(2001).			
RN	[2]			
RP	SEQUENCE OF 624-628; 676-682; 938-948; 1156-1168; 1176-1185;			
RP	1199-1205 AND 1217-1224.			
RC	STRAIN=Isolate 1061;			
RX	MEDLINE=96270743; PubMed=8665912;			
RA	Hilden P., Savolainen K., Tynnelae J., Vuento M., Kuusela P.;			
RT	"Purification and characterization of a plasmin-sensitive surface			
RT	protein of Staphylococcus aureus.";			
RL	Eur. J. Biochem. 236:904-910(1996).			
CC	-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AF115379; AAD09131.1; -.			
DR	InterPro; IPR005877; Gpos_Y5IRK.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF04650; Y5IRK signal; 1.			
DR	TIGRFAMS; TIGR01167; LPXTG anchor; 1.			
DR	TIGRFAMS; TIGR01168; Y5IRK signal; 1.			
DR	PROSITE; PSS00847; GRAM_POS_ANCHORING; 1.			
KW	Antibiotic resistance; Glycoprotein; Cell wall; Peptidoglycan-anchor;			
FT	Repeat; Signal.			
FT	SIGNAL 1 48			
FT	CHAIN 49 1601			
FT	PROPEP 1602 1637			
FT	DOMAIN 1301 1582			
FT	POTENTIAL.			
FT	METHICILLIN-RESISTANT SURFACE PROTEIN.			
FT	REMOVED BY SORTASE (POTENTIAL).			
FT	141 X 2 AA TANDEM REPEATS OF D-[SAG].			

RA Ritchie H.H.;
 RT "Human dentin phosphophoryn nucleotide and amino acid sequence.";
 RL Eur. J. Oral Sci. 106:1043-1047(1998).
 RN [3]
 RP DISEASE.
 RX MEDLINE=21096971; PubMed=11175779;
 RA Zhang X., Zhao J., Li C., Gao S., Qiu C., Liu P., Wu G., Qiang B.,
 Lo W.H.Y., Shen Y.;
 RT "DSPP mutation in dentinogenesis imperfecta Shields type II.";
 RL Nat. Genet. 27:151-152(2001).
 RN [4]
 RP VARIANTS DFNA39/DG11 THR-17 AND PHE-18.
 RX MEDLINE=21096982; PubMed=11175790;
 RA Xiao S., Yu C., Chou X., Yuan W., Wang Y., Bu L., Fu G., Qian M.,
 Yang J., Shi Y., Hu L., Han B., Huang W., Liu J., Chen Z.,
 Zhao G., Kong X.;
 RT "Dentinogenesis imperfecta 1 with or without progressive hearing loss.
 RA is associated with distinct mutations in DSPP.";
 RL Nat. Genet. 27:201-204(2001).
 CC -I- FUNCTION: DSP may be an important factor in dentinogenesis. DPP
 CC may bind high amount of calcium and facilitate initial
 CC mineralization of dentin matrix collagen as well as regulate the
 CC size and shape of the crystals.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- TISSUE SPECIFICITY: Expressed in teeth. DPP is synthesized by
 CC odontoblast and transiently expressed by pre-ameloblasts.
 CC -I- PTM: DSP is glycosylated.
 CC -I- DISEASE: Defects in DSPP are the cause of dentinogenesis
 CC imperfecta 1 (DG11) [MIM:125490]; also known as dentinogenesis
 CC imperfecta Shields type II (DGI-II). DG11 is an autosomal dominant
 CC disorder in which both the primary and the permanent teeth are
 CC affected. It occurs with an incidence of 1:8000 live births. The
 CC teeth are amber and opalescent, the pulp chamber being obliterated
 CC by abnormal dentin. The enamel, although unaffected, tends to
 CC fracture, which makes dentin undergo rapid attrition, leading to
 CC shortening of the teeth.
 CC -I- DISEASE: Defects in DSPP are the cause of autosomal dominant
 CC deafness type 39 with dentinogenesis imperfecta 1 syndrome
 CC (DFNA39/DG11) [MIM:605594]. Affected individuals present DG11
 CC associated with early onset progressive sensorineural high-
 CC frequency hearing loss.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF163151; AAF42472.1; .
 DR EMBL; AF094508; AAD16120.1; .
 DR Genbank; HGNC:3054; DSPP.
 DR MTM; 125485; .
 DR MTM; 125490; .
 DR MTM; 605594; .
 KW Biomeralization; Extracellular matrix; Signal; Glycoprotein;
 KW Phosphorylation; Stalic acid; Disease mutation; Deafness.
 FT SIGNAL 1 15
 FT CHAIN 16 1253
 FT CHAIN 16 462
 FT CHAIN 463 1253
 FT DOMAIN 439 1253
 FT SITE 488 490
 FT MOD_RES 259 259
 FT CARBOHYD 41 41
 FT CARBOHYD 49 49
 FT CARBOHYD 81 81
 FT CARBOHYD 130 130
 FT CARBOHYD 150 150
 FT CARBOHYD 190 190
 FT CARBOHYD 191 191
 FT CARBOHYD 209 209

FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 17 17 P -> T (in DFNA39/DG11).
 FT FTId=VAR 012280.
 FT VARIANT 18 18 V -> P (in DFNA39/DG11).
 FT FTId=VAR 012281.
 FT CONFLICT 685 690 MISSING (IN REF. 2).
 FT CONFLICT 744 744 S -> SNSSDS (IN REF. 2).
 FT CONFLICT 799 799 N -> D (IN REF. 2).
 FT CONFLICT 836 836 S -> C (IN REF. 2).
 FT CONFLICT 850 850 S -> G (IN REF. 2).
 FT CONFLICT 888 888 MISSING (IN REF. 2).
 FT CONFLICT 963 963 G -> S (IN REF. 2).
 FT CONFLICT 1005 1005 N -> D (IN REF. 2).
 FT CONFLICT 1025 1025 S -> G (IN REF. 2).
 FT CONFLICT 1047 1047 N -> D (IN REF. 2).
 FT CONFLICT 1053 1053 D -> N (IN REF. 2).
 FT CONFLICT 1065 1065 G -> D (IN REF. 2).
 FT CONFLICT 1080 1080 D -> E (IN REF. 2).
 FT CONFLICT 1095 1095 E -> D (IN REF. 2).
 FT CONFLICT 1101 1101 D -> E (IN REF. 2).
 FT CONFLICT 1104 1104 D -> N (IN REF. 2).
 FT CONFLICT 1132 1132 S -> R (IN REF. 2).
 SQ SEQUENCE 1253 AA; 126426 MW; DAE240653904ED4A CRC64;
 Query Match 12.5%; Score 708.5; DB 1; Length 1253;
 Best Local Similarity 23.4%; Pred. No. 3.8e-20;
 Matches 241; Conservative 147; Mismatches 405; Indels 239; Gaps 23;
 QY 44 LGHNEAKAEENS-----VQVQDSNTDELSDNSDSSDEKDVNNNSQINT 92
 DB 436 VGHSTGTSGSDSDGYDSDYDFDKSMQGDPSNSDNGNDANDNESDSSRSGDASYS 495
 QY 93 DNNNQI1KKEETNYD--GIEKRSDETRSTTV-----DNEATFLKOTPDQNT 140
 DB 496 DES-----KDNNGSDSKGAEDDSDSTDTNNSDNGNNGNDNDKSDGKGSDS- 549
 QY 141 HITEEEVKSSSVESSESSSIDTAQAPSHTTINREESVQTSNVEDSHVSDFANSKIKES 200
 DB 550 --SDSDSDSSSDS 607
 QY 201 TSGKRENTIEOPKVKVERSTTSPQSGYTNIDKISNQDELLMLPINEYENKARPISTTS 260
 DB 608 SDSDSDSDSDSDSDS-KSDSKSE-SDSDSDSKSDSDS-----NSDSDSDSDSDS 659
 QY 261 AOPSIKRVTVNQLAABQGSNNVHLIKVTDQSITEGYDSDSEGVKAKDAENLIYDVTFEVD 320
 DB 660 SNS-----NSDS 708
 QY 321 DKYKSGDTMTVD-IDKNTVPSDLTDSFTPIKIDNSGEIATGYDNKNKQIYTFDYV 379
 DB 709 DSSDS 768
 QY 380 DKYENTKAHLKLTYSYDKSKVPNNNTKL DVEYKLTALSSVKNKTIYVEQRPNEHTANLQS 439
 DB 769 DSSDS-----SDSDSDSDSDSDS-----NDSNSSDSDS 795
 QY 440 MFTNIDTKHVTQTIYINPLRYSAKETNVNLSNGNDEGSTIIDSTIIKVKYGVGNQL 499
 DB 796 DSSNSSDSDSDSDSDS-----DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 838
 QY 500 PDSNRIYDSEYEDVTNDDYQAQGNNDVNIINFGNIDSPYIIKVISKIDPNKDDYTTIQ 559
 DB 839 SDS 866
 QY 560 TVTMTQTTINEYTGFBFTASVDNTAFSTSGGQGGDLPPEKTYKIGDYVWEDVKDGIQN 619
 DB 867 -----SDS 885
 QY 620 TNDNEKPLSNVLVLTYPDGTSKSVTDEGKYQFDGLKNGLTYKITFETPEGYTPTLKH 679

Db 886 SDSNSSSS-----DSSDSNSSSDSS-----DS 911
 Qy 680 SGTNPALDSEGNVYVINGQDDMTIDSGFYQTPKYSLGNTGVYVDTNKGIGQGDKEGIS 739
 Db 912 SNSSDSDSDSDSSSDSSSSSSSDSSSS-----DSSNSSSDSDSDSSSDSSN 955
 Qy 740 GVKYTLKDENGNIIITTTTDENGKYQFNLNSGNYVHFDPKPSGWTQTTTSGDDDDQDA 799
 Db 956 SSSDSNGGSSN--SSDSNSSSDSDSDSSSD-----SSDSNSSSDSDSDSSN 1005
 Qy 800 DGEVHVITDHDFFSDNGVYVDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 859
 Db 1006 SSSDSSDSG 1065
 Qy 860 DS 919
 Db 1066 SSSDSSD 1125
 Qy 920 DS 979
 Db 1126 SSSDSSDSNE 1185
 Qy 980 DS 1039
 Db 1186 SSSDSSDSE--- 1242
 Qy 1040 DLGNSDKSTKD 1051
 Db 1243 --GSDSNHSTSD 1252

RESULT 3

ID DSGP MOUSE STANDARD; PRT; 934 AA.
 AC P97399; 070567;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dentin sialoprotein precursor (Dentin matrix protein-3) (DMP-
 DE 3) [Contains: Dentin phosphoprotein (Dentin phosphophoryn) (DPP);
 DE Dentin sialoprotein (DSP)].
 GN DSGP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Molar;
 RX MEDLINE=97150835; PubMed=8995371;
 RA MacDougall M., Simmons D., Luan X., Nydegger J., Feng J.Q., Gu T.T.;
 RT "Dentin phosphoprotein and dentin sialoprotein are cleavage products
 expressed from a single transcript coded by a gene on human chromosome
 4. Dentin phosphoprotein DNA sequence determination.";
 RL J. Biol. Chem. 272:835-842 (1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND REVISIONS TO C-TERMINUS.
 RC STRAIN=129/SvJ; TISSUE=Liver;
 RX MEDLINE=98211969; PubMed=9545272;
 RA Feng J.Q., Luan X., Wallace J., Jing D., Ohshima T., Kulkarni A.B.,
 RA D'Souza R.N., Kozak C.A., MacDougall M.;
 RT "Genomic organization, chromosomal mapping, and promoter analysis of
 the mouse dentin sialoprotein (dspp) gene, which codes for both
 dentin sialoprotein and dentin phosphoprotein.";
 RL J. Biol. Chem. 273:9457-9464 (1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ; TISSUE=Liver;
 RA Sfeir C., Butler S., Lin E., George A., Veis A.;
 RT "From mouse to zebrafish-dentin matrix proteins genomic
 characterization.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP TISSUE SPECIFICITY.
 RX MEDLINE=98055479; PubMed=9395101;
 RA Ritchie H.H., Berry J.E., Somerman M.J., Hanks C.T., Bronckers A.L.,
 RA Hotton D., Papagerakis P., Berdal A., Butler W.T.;
 RT "Dentin sialoprotein (DSP) transcripts: developmentally-sustained
 expression in odontoblasts and transient expression in
 pre-odontoblasts.";
 RL Eur. J. Oral Sci. 105:405-413 (1997).
 RN [5]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21096982; PubMed=11175790;
 RA Xiao S., Yu C., Chou X., Yuan W., Wang Y., Bu L., Fu G., Qian M.,
 RA Yang J., Shi Y., Hu L., Han B., Wang Z., Huang W., Liu J., Chen Z.,
 RA Zhao G., Kong X.;
 RT "Dentinogenesis imperfecta 1 with or without progressive hearing loss
 is associated with distinct mutations in DSP.";
 RL Nat. Genet. 27:201-204 (2001).
 CC -I- FUNCTION: DSP may be an important factor in dentinogenesis. DPP
 may bind high amount of calcium and facilitate initial
 mineralization of dentin matrix collagen as well as regulate the
 size and shape of the crystals.
 CC -I- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -I- TISSUE SPECIFICITY: Expressed in teeth, mainly in odontoblasts and
 transiently in pre-odontoblasts. Found in the inner ear.
 CC -I- PTM: DSP is glycosylated.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U67916; AAC12787.1; -
 DR EMBL; A002141; CAA05208.1; ALT_SEQ.
 DR EMBL; AF135799; AAD42781.1; ALT_SEQ.
 DR MGI; 109172; Dppp.
 KW Biomineralization; Extracellular matrix; Signal; Glycoprotein;
 KW Phosphorylation; Sialic acid.
 FT SIGNAL 1 17
 FT CHAIN 18 934
 FT CHAIN 45 451
 FT CHAIN 452 934
 FT DOMAIN 419 934
 FT SITE 479 481
 FT MOD_RES 227 227
 FT MOD_RES 254 254
 FT MOD_RES 279 279
 FT MOD_RES 293 293
 FT MOD_RES 299 299
 FT MOD_RES 314 314
 FT MOD_RES 336 336
 FT MOD_RES 349 349
 FT CARBOHYD 54 54
 FT CARBOHYD 84 84
 FT CARBOHYD 130 130
 FT CARBOHYD 190 190
 FT CARBOHYD 313 313
 FT CARBOHYD 373 373
 SQ SEQUENCE 934 AA; 93901 MW; A618789D8A57249A CRC64;
 Query Match 12.0%; Score 675.5; DB 1; Length 934;
 Best Local Similarity 26.2%; Pred. No. 4.8e-19;
 Matches: 279; Conservative 107; Mismatches 439; Indels 241; Gaps 39;
 QY 49 AKAENSVD-VKDSNTDDELSDSDQ--SSDEKND-----VINNN 87
 DB 46 AAQNELSINSTNSNSPDGSEIGQVLSGDKYKRGNGSESITHVGKDPFPTQILYNE 105
 QY 88 QSINTDNNQIKKEETNNYDGIEKRSBDRTESTTVDENESTFLOKTPQDNTLHTEEV 147
 DB 106 QGNTAEHNDI-----ETGHDGVHARGN---STANGRSQGVIV-----ENAEAESSV 153


```
QY 148 K-ESSSVESNSSIDTAQPSHTTINREE-----SVQTSNVDERS-HVSPFANSKIKESWT 201
DB 154 HGQAQNTKSGGASDVSONGATLVQENEPPEASIRKSTNHEAGIRGSGVATHTTTPQRE 213
QY 202 ESCKEENTTEOPNKVEDS---TTQPSGY-TNIDEKISNQDELLNPLINEYENKARPL 256
DB 214 GLGSENGTEVTPSIGEDAGLDTDGSPGNGVEDEDTGSGD-----256
QY 257 STTSAQPSIKRV-----TVNQLAAEQSNVNHLIKVTQDSITEGYDPS---EGVIKAHADAE 309
DB 257 -GEGAAGDGRSHDGTGQGGQSHGCGNTDHR---GQSSVSTEDDSDSKQEGFPNGHNGD 312
QY 310 NLIYDVTFFVDVKVSGDTWTVIDIKNTV-PSLTD-----STPIKIKNSGEITA 360
DB 313 N-----SSEENGVEEGDSTQATQDKELKSPKDRDAEGGIIISQSBACPSGKSDQGIET 366
QY 361 TGYDNK-NKOITVYTFDYDKYENIKAKHLKLTYSIDKSKVPNNNTKLDEYKVTALSSVN 419
DB 367 EG--PNKGNKSI-----TKESGKSGKSDSGHGVLDKNS-----403
QY 420 KTIIVYQRPENRNTANLQSMFTNIDTKNHTVBQTIYINPLRYSAKETNVNISNGDEGS 479
DB 404 -----PKQGES-----DKPQGTAEK-----SAAHNLGHSRIGSSN 435
QY 480 TIIDSTIIKVKGNQNLPSNRIDYDYSEYEDVTDNDYQALGNNDVNIINFGNIDSPY 539
DB 436 SDGHDS-----YFDDDESQGDGPKSDSGSDSGSDST-----NSESANESGSDGASY 484
QY 540 IIKVISKYDKNDDYTTIQTQVTMTQITNEYTGFRTASVDNTIAFTSFGQGGQLPPE 599
DB 485 TSDESDDDNDSDH-----AGEDD-----SDSDSGDGDSD-----516
QY 600 KTYKIGDYVWEDVDKQGIQNT-NDNEKPLSNVLTLTPDGTGSKSVRTD---EDGKYQFDG 656
DB 517 ---SNGDGSSEDEKDESDSDHNS-----DSEKSDSDSDSDSDSDSDSDSDS 561
QY 657 LKNGLYKTFETPEGYPTLKHSGTNPALDSGNSVWVTINGQDQMTTDSGYQYQPKYS 716
DB 562 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 604
QY 717 LGNVVWYDT-----NKDGIQGDKEKISGVKVKTLKDNNGNIISTTTDENGKQFQDN 768
DB 605 -----DTCSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 657
QY 769 LNSGNYIVHFKPSGMTQTTTDSGDDEQDADGEEVHVTTITDHDFFSIDNGYYDDSDSD 828
DB 658 SDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 709
QY 829 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
DB 710 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 768
QY 889 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 948
DB 769 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 828
QY 949 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1008
DB 829 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
QY 1009 SDSDSDSDSDSDSDG---SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1051
DB 889 SKDSDSDSDGSDSGNSGNSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 933
```

RESULT 4

DSDP RAT

ID DSDP RAT STANDARD; PRT; 687 AA.

AC Q62598; P70578; Q9R057;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

```
DE Dentin sialophosphoprotein precursor (Contains: Dentin phosphoprotein
DE (Dentin phosphophoryn) (DPP); Dentin sialoprotein (DSP)).
GN DSDP OR RDSF2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM DPP-2).
RC STRAIN=Sprague-Dawley;
RA MEDLINE=20435277; PubMed=10978503;
RX Ritchie H.H., Wang L.-H.;
RT "The presence of multiple rat DSP-PP transcripts.";
RL Biochim. Biophys. Acta 1493:27-32(2000).
RN [2]
RP SEQUENCE OF 1-387 FROM N.A., AND SEQUENCE OF 18-26.
RC STRAIN=Sprague-Dawley; TISSUE=Odontoblast;
RX MEDLINE=94148875; PubMed=8106414;
RA Ritchie H.H., Hou H., Veis A., Butler W.T.;
RT "Cloning and sequence determination of rat dentin sialoprotein, a
RT novel dentin protein.";
RL J. Biol. Chem. 269:3698-3702(1994).
RN [3]
RP SEQUENCE OF 421-687 FROM N.A. (ISOFORM DPP-1).
RC STRAIN=Sprague-Dawley;
RX MEDLINE=96355551; PubMed=8702961;
RA Ritchie H.H., Wang L.-H.;
RT "Sequence determination of an extremely acidic rat dentin
RT phosphoprotein.";
RL J. Biol. Chem. 271:21695-21698(1996).
RN [4]
RP SEQUENCE OF 1-19 FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=99333695; PubMed=10403786;
RA Yamazaki H., Kunisada T., Miyamoto A., Tagaya H., Hayashi S.-I.;
RT "Tooth-specific expression conferred by the regulatory sequences of
RT rat dentin sialoprotein gene in transgenic mice";
RL Biochem. Biophys. Res. Commun. 260:433-440(1999).
RN [5]
RP SEQUENCE OF 29-33; 70-79; 93-109; 136-148; 162-188; 266-308; 398-423
RP AND 426-438, AND PHOSPHORYLATION OF SER-292 AND SER-298.
RX MEDLINE=21125612; PubMed=11042175;
RA Qin C., Cook R.G., Orkiszewski R.S., Butler W.T.;
RT "Identification and characterization of the carboxyl-terminal region
RT of rat dentin sialoprotein.";
RL J. Biol. Chem. 276:904-909(2001).
RN [6]
RP TISSUE SPECIFICITY.
RX MEDLINE=98055479; PubMed=9395101;
RA Ritchie H.H., Berry J.E., Somerman M.J., Hanks C.T., Bronckers A.L.,
RA Hottel D., Papagerakis P., Berdal A., Butler W.T.;
RT "Dentin sialoprotein (DSP) transcripts: developmentally-sustained
RT expression in odontoblasts and transient expression in
RT pre-ameloblasts.";
RL Eur. J. Oral Sci. 105:405-413(1997).
CC -!- FUNCTION: DSP may be an important factor in dentinogenesis. DPP
CC may bind high amount of calcium and facilitate initial
CC mineralization of dentin matrix collagen as well as regulate the
CC size and shape of the crystals.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=DPP-1; Synonyms=PP240;
CC IsoId=Q62598-1; Sequence=Displayed;
CC Name=DPP-2; Synonyms=PP171;
CC IsoId=Q62598-2; Sequence=VSP_003855;
CC Note=No experimental confirmation available;
CC -!- TISSUE Specificity: Specifically expressed in teeth, mainly in
CC odontoblasts and transiently in pre-ameloblasts.
CC -!- PTM: DSP is glycosylated.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 380.
CC -----
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF247187; AAK96895.1; -
EMBL; U02074; AAA18932.1; ALT_FRAME.
EMBL; U63111; AAC52774.1; -
EMBL; AF114987; AUA48588.1; ALT_TERM.
KW Biminerization; Extracellular matrix; Signal; Glycoprotein;
Phosphorylation; Sialic acid; Alternative splicing.
FT SIGNAL 1 17
FT CHAIN 18 687 DENTIN SIALOPHOSPHOPROTEIN.
FT CHAIN 18 447 DENTIN SIALOPHOSPHOPROTEIN.
FT CHAIN 448 687 DENTIN SIALOPHOSPHOPROTEIN.
FT MOD RES 57 226 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD RES 226 226 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD RES 253 253 PHOSPHORYLATION (BY CK1) (POTENTIAL).
FT MOD RES 278 278 PHOSPHORYLATION (BY CK1) (POTENTIAL).
FT MOD RES 292 292 PHOSPHORYLATION (BY CK2).
FT MOD RES 298 298 PHOSPHORYLATION (BY CK2).
FT MOD RES 315 315 PHOSPHORYLATION (BY CK1) (POTENTIAL).
FT MOD RES 319 319 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD RES 329 329 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD RES 337 337 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD RES 345 345 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD RES 366 366 PHOSPHORYLATION (BY CK1) (POTENTIAL).
FT CARBOHYD 55 55 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 567 635 Missing (in isoform DPP-2).
FT CONFLICT 74 74 N -> D (IN REF. 5; AA SEQUENCE).
FT CONFLICT 564 564 S -> T (IN REF. 3).
SQ SEQUENCE 687 AA; 70179 MW; 9A845EED6AA31B63 CRC64;

Query Match 10.78; Score 606.5; DB 1; Length 687;

Best Local Similarity 32.08; Pred. No. 1.4e-16;
Matches 211; Conservative 68; Mismatches 250; Indels 131; Gaps 30;

QY 453 QTVINPLRYSAKET-----NWNISGNGEGST-IIDSTIIKVKY-GDN 496
DB 97 QPILANAQNSAKEREVETVYHGDHAGGENSTANGIRQGVIAEAEKESKVGQP 156
QY 497 QNLPDSNRIYDYGEYEDVT---NDVYALGNNDVNINFGNIDSPYIIKISKYDPNKDD 553
DB 157 HQDTKTGLASDTSGQDATLVQENPQVAGSKNSTNEVTHGSGV---AAQETTPQREG 213
QY 554 YTIQIVTWTQVINEYTGFRASVDNTAFSTSSGQGGDLPPEKTYKIGYVWEDV- 612
DB 214 EGENOQAFVTPSIGAG-----LDNT-----EGSPSGNGIEBEDTGS-GDGVGADAG 262
QY 613 -----DKDGIQ-----NTNDNEKPLSNVLVTITPDGTSKSVRTDEDGKYQFDGLKNG 660
DB 263 DGRESHDTGTEGSCQSGGNDR-----CQGSVSTEDDDSKQESGPNQ 307
QY 661 LTYKITPETPEGYPTTLKHSHTNAPLD-----SEGNSVWYTINGQDDMTIDSGFYTPKY 715
DB 308 -----RGCDNTSSSEETGIEEGDGTQTQDNQNLSPTEGGIISQAEA 349
QY 716 SLGNVYVYDNTKDIQGDDEKGLS-GVKVTLKDBNGNIIITTTDENGKY--QFDNLNSG 772
DB 350 C-----PSGQSONQLETEGSGTGNKSIITKESGKL--SGSKDNGHGMELDKRNSP 400
QY 773 NYIVHFDKPSGMTQ---TTTDSGD---DEQDADGEVHVHTIITDHDFFSIDNGYDDESD 826
DB 401 KQ-GESEKPOGAEEKSDTHNMGMHSRIGSSNSDGHDSY-----DFDESQOG---DDPNS 452

QY 827 SDSDSDSDSDSDSDS-----DSDSDSDSDSDSDSDSDSDSDSDSDS 874
DB 453 SDSNSGSDGSDANSAIENGHGDAVTSDESSNGSDSDSHAGSDSDTSDTD-D 511
QY 875 SDSDSDSDSDS-DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 933
DB 512 SDNSGDDDESXKDKSDNSNHDNDSSEKSDS-SDSDSDSDSDSDSDSDSDSDS 570
QY 934 DSNSDSDSDSDSDSDSDS-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 992
DB 571 SDSSDTSD 639
QY 993 SDSDSDSDSDSDS-DSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1051
DB 630 DSNSDSDSDSDSDSDSKSDSTSDSDSDSDSDSDSDSDSDSDSDSDSDSD 686
RESULT 5
SR40 YEAST STANDARD; PRT; 406 AA.
ID SR40 YEAST
AC P32583;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Suppressor protein SRP40.
GN SRP40 OR YKR092C OR YKR412A.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RC STRAIN=ATCC 28383 / FL100;
RA Lalo D., Carles C., Sentenac A., Thuriaux P.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94205265; PubMed=8154186;
RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT The complete sequence of a 15,820 bp segment of Saccharomyces
RT cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three
RT new open reading frames."
RL Yeast 9:1349-1354(1993).
CC -!- FUNCTION: Not known; weak suppressor of a mutant of the
CC subunit AC40 of DNA dependant RNA polymerase I and III.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L11275; AAA35091.1; -
DB EMBL; X73541; CAAS1946.1; -
DB EMBL; Z28317; CAAB2171.1; -
DB PIR; S38170; S38170.
DB Germonline; 140071; -
DB SGD; S0001800; SRP40.
DB GO; GO:0005730; C:nucleolus; IDA.
DB InterPro; IPR007718; SRP40_C.
DB Pfam; PF05022; SRP40_C; 1.
FT DOMAIN 25 374 ASP/SER-RICH
FT CONFLICT 400 400 G -> N (IN REF. 1).
SQ SEQUENCE 406 AA; 41015 MW; 8EA007695AF4BA1D CRC64;

Query Match 9.8%; Score 554; DB 1; Length 406;
Best Local Similarity 46.8%; Pred. No. 7.5e-15;
Matches 137; Conservative 25; Mismatches 93; Indels 38; Gaps 6;


```
Db 4005 RYNTAMTVLERSRIYYTNTSRGMYILSTILSLAKNGFCSPQPPSEVDDKNLQEG--- 4061
Qy 712 TPKYSLGNYVYDTNKDGIQGDDEKISGVKVTLKDENGNIISTTTTDENGKYQFNLS 771
Db 4062 ---TGLGDGSGCAQNNKDVQDEB-----LTEDAQEN-----KEQQDKDERDENE 4105
Qy 772 GNYIVHDFKPSGWTQTITSDGDDDEQADGEEVHV-----TITDHDDFSIDNGYDDESD 826
Db 4106 DDAVEEGMDWAGELEDLSNGEENDEDTSEEBELDEIDDLNEDDPNAIDDKQWDDKA- 4164
Qy 827 SDSDSDSDSDSDS-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 885
Db 4165 SDNSKEKOTDNLGDKNQEEVQAAENDEQORDKNGEGEDPNAPEDGEEIENDENAE 4224
Qy 886 DSNSDSDSDSDSDSDSDSDSDSDSDSDSDSDS-SDSDSDSDSDSDSDSDSDSDSDSD 944
Db 4225 ENDVGEQDEVKDEEGE-DLEAVNPETITLDPEDMNLDSHESBEDVDMSCMPDDL 4283
Qy 945 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1004
Db 4284 KEEVGNDEEVKQESGIESNENDEPGFEEDAGETETALDEEGABEDVDMTDECKE 4343
Qy 1005 SDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1064
Db 4344 ENGPEEQAMSDRELKQDAAMENKEKGCGQNTGLDGVEEKADTEDIDQEAQVQDSGS 4403
Qy 1065 KG 1066
Db 4404 KG 4405

RESULT 13
VG48 HSVSA STANDARD; PRT; 797 AA.
AC Q01033;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Hypothetical gene 48 protein.
GN 48 OR EDL45.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333689; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -I- SIMILARITY: TO EBV BRP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64346; CAA45671.1; -.
DR InterPro; IPR008550; DUF832.
DR Pfam; PF05734; DUF832; 1.
KW Hypothetical protein.
SQ SEQUENCE 797 AA; 89867 MW; 9E294234AD950E23 CRC64;

Query Match 5.9%; Score 333; DB 1; Length 797;
Best Local Similarity 18.8%; Pred. No. 3.9e-06;
Matches 116; Conservative 104; Mismatches 275; Indels 122; Gaps 22;

Qy 506 YDSEYEDVTNDYQAQGNNDNVNFGNIDSPYIIKVISKDPNKDDYTTIQVTMQT 565
```

```
Db 194 YDFNEHK-----NEISLVATCINCCWLYFMLQY--MSDDLAEALN-KT 237
Qy 566 TINEYTGFEPTASYDNTIAFSTSSGQG-----GDLPPKTYKIGYVWEDVDKGI 617
Db 238 YLALHPND--KASYSNILKFLTNSHREHVTKVNVKAPMQSSLYKI-----I 283
Qy 618 QNTNDNEKPSLVV-----TLTYP--DGTSKSVRTDEGKYQFDGL 657
Db 284 KQTEKNPSPKTKLMIILSGRIGMDLFCOSVLKAPLIDHKLSVSEYED----FD-- 337
Qy 658 KXGLTVKIFETPEGYTPTLKHSNTNPALDSGNSV-----WVTINGQD----- 701
Db 338 EREVELICSDDEV-----SEDGNLCVLDDESESVNSVALRQVLTVDQANEKEYKII 391
Qy 702 DMTIPSGFYQTPKYSIGNVYVDTNKDGIQGDDEKISGVKVTLKDENGNIISTTTT 761
Db 392 DKSDDRDDRDKDEYLENEE--YNRDEEDEDGEDEKDEKEGEDEGD--GEDEGEDE 448
Qy 762 GKYQFNLANSYIVHDFKPSGWTQTITSDGDDDEQADGEEVHV-VTITDHDDFSIDNGY 820
Db 449 GEDEGDEGEG-----DEGEDEGEDEDEDEGEDEGEDEGEDEGD 491
Qy 821 YDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 880
Db 492 EGDGEDEGD--EGDEGKDEGDEGKDEGDEGDEGEDEGEDEGEDEGEDEGEDEG 550
Qy 881 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 940
Db 551 EDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEG 610
Qy 941 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1000
Db 611 EGDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 670
Qy 1001 SDSDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060
Db 671 GDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 725
Qy 1061 -----DYGS---KGTLL 1068
Db 726 VVQNPFYNNLQKSTL 742

RESULT 14
ALAI CANAL STANDARD; PRT; 1419 AA.
ID ALAI CANAL
AC O13368;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Agglutinin-like protein ALAI precursor (Agglutinin-like adhesin).
GN ALAI OR ALG5.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98053977; PubMed=9393828;
RA Gaur N.K., Klotz S.A.;
RT "Expression, cloning, and characterization of a Candida albicans
RT gene, ALAI, that confers adherence properties upon Saccharomyces
RT cerevisiae for extracellular matrix proteins.";
RL Infect. Immun. 65:5289-5294(1997).
CC -I- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -I- PTM: N-glycosylated and O-glycosylated (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```


FT SIGNAL 1 20
FT CHAIN 21 913
FT PROPEP 914 937
FT TRANSMEM 42 62
FT DOMAIN 332 576
FT DOMAIN 569 576
FT DOMAIN 577 814
FT DOMAIN 610 753
FT REPEAT 610 613
FT REPEAT 666 669
FT REPEAT 680 683
FT REPEAT 690 693
FT REPEAT 698 701
FT REPEAT 738 741
FT REPEAT 750 753
FT DOMAIN 901 904
FT LIPID 913 913

FT CARBOHYD 16 16
FT CARBOHYD 236 236
FT CARBOHYD 449 449
FT CARBOHYD 488 488
FT CARBOHYD 580 580
FT CARBOHYD 585 585
FT CARBOHYD 595 595
FT CARBOHYD 603 603
FT CARBOHYD 619 619
FT CARBOHYD 631 631
FT CARBOHYD 641 641
FT CARBOHYD 649 649
FT CARBOHYD 711 711
FT CARBOHYD 747 747
FT CARBOHYD 759 759
FT CARBOHYD 773 773
FT CARBOHYD 897 897
FT CARBOHYD 913 913
SQ SEQUENCE 937 AA; 93699 MW; 178021215BA926 CRC64;

Query Match 5.6%; Score 315; DB 1; Length 937;

Best Local Similarity 21.5%; Pred. No. 2.3e-05;

Matches 189; Conservative 133; Mismatches 349; Indels 208; Gaps 30;

QY 317 FEYDDKVKSGDTWVIDKNTVPSDLTDSFTPK-----IKNSGRIIA----- 360
DB 36 FHGDVVKHSGATWAI---LGTTLCSPPFGGLEVEKGASLFIKSDNGPVLALNVALSTLVRP 92
QY 361 ---TGTVDNKNQKIT-----YFTDYVDKYNKIKHLKLTYSIDKSKVNNNTKLD 408
DB 93 VINNGVISLNSKSTSFNSFDIGCSFTNGEYILDSSGLVKSTALYAREWTNG----- 148
QY 409 VEYKTALSSVKNKTIYEQRPENRNTANLQSMFTNIDTKNHTVEQTIYINP---LRYSAK 465
DB 149 -----LIVAYQ--NQKAAGNIAF-----GTAYQITNNGQICLRHODF 184
QY 466 ETWNVINGN-----DE-----GSTII-----DDSTIKVKYKVDNQ----- 497
DB 185 VPATKINGTCVTADEDTWIKLNTILSVEPTHNFYKLSKSLIVHAVSSNQTFTHVGF 244
QY 498 -----NLP-DSNRIDYSEYEDVTN-----DDYAOLGNNDVNINFGNIDSPY 539
DB 245 GNGNKLGLTLPLTGNRDHFRFEYPTDGIQLRADALPQYFKIKGYDSKL-FRIVNSRG 303
QY 540 IIKVISKYD---PNKDDYTTIQOTVTMQTTINBYTGFEFTASYDNTIAFSTSSGQGGDL 596
DB 304 LKNAVTVYDGPVNNIEPAVCLIPCTNGPSAPESDNLPTTSSITSSYSSAATESSV 362
QY 597 PPEKTKIGDYVEDVDKGIQNTN-DNEKPLNVLTLTPDGTSKSVRTDEDEGKQPD 655
DB 363 VSESSSAV-----DSLTSLSLSKSESDVWSSTTNISSSTAIETTNSESSTD 412
QY 656 GLKNGLTGYKITPTEGTYPTLKHSGTNPALDSEGNVWVTINGQDDMTIDSGFYQTPKY 715
DB 413 AGSSSIS-----QSESSSITSSSETS-----SSESMSASSTTASNTSIETDGIIVSQS--- 462

Search completed: October 5, 2004, 20:49:06
Job time : 32.8315 secs

QY 716 SLGNVYVYDTNKDGIQGDDEKIGISGVKVLKDBENGNIISITTTDENGKYQFDNLN----- 770
DB 463 -----BSSGNALSSTEQSITSSPGQSTIYVNSTVTITSCDENKCTEDVVITFTTV 514
QY 771 -----SGNY-----IVHFKPS-----GMTQTTT 789
DB 515 PCSTDCVPTTGIPMSTSYTQRTVTSITINCDEVSCQDVVYTTNVPHTTVDATTTTT 574
QY 790 DSGDDDEQDADGEEVHVITITDHDPSIDNGYYDDESD-SDSDSDSDSDSDSDSDSDSDSDSD 848
DB 575 STGDNLS-----TGGNESGSHGPGNGSTEGSGNGSGAGSNEGSGPNNGSG 622
QY 849 SD 908
DB 623 SGSEGGSNNGSGSDSGNNGSGSGSTEGSEGGSGNEGSGSGSQPGPNEG 682
QY 909 SD 968
DB 683 SEGGSGSNNGSHGNEGSGSGSGNNGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSN 742
QY 969 SD 1028
DB 743 PGAGNGSNEGSGGSGNGSEAGSGGSGGPNNGSGSHNDGSGSGSHNQGNPGAGSGSGSE 802
QY 1029 SD 1061
DB 803 SGSKAGSHSGNEGAKTDSIEGFHTESK-PGFNTGAHTD 840

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:42:37 ; Search time 39.7921 Seconds
(without alignments)
2639.748 Million cells updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKKNLLTKKPIANKSN.....FAGLGALLLKRRKRNKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES											
Result No.	Score	Query Match	Length	DB	ID	Description					
1	5646	100.0	1092	2	T30214	fibrinogen-binding					
2	2698.0	47.8	1166	2	T28680	fibrinogen-binding					
3	2589	45.9	1141	2	E89824	hypothetical prote					
4	2234	39.6	1385	2	D89824	hypothetical prote					
5	2093.5	37.1	953	2	C89824	hypothetical prote					
6	1880	33.3	1315	2	T28679	fibrinogen-binding					
7	1652	29.3	989	2	D89852	fibrinogen-binding					
8	1628	28.8	933	2	S41539	fibrinogen-binding					
9	1536	27.2	877	2	F90070	Clumping factor B					
10	785	13.9	882	2	AG1671	probable peptidogl					
11	782.5	13.9	903	2	AG1299	probable peptidogl					
12	624.5	11.1	2271	2	F90073	hypothetical prote					
13	554	9.8	406	2	S38170	SRP40 protein - ye					
14	554	9.8	3394	2	T18501	hypothetical prote					
15	549	9.7	2570	2	T17451	fimbriae-associate					
16	525.5	9.3	1038	2	H90053	hypothetical prote					
17	520	9.2	940	2	S19702	fibronectin-bindin					
18	502.5	8.9	961	2	G90053	hypothetical prote					
19	487.5	8.6	1018	2	A32192	fibronectin-bindin					
20	471	8.3	334	2	A54138	acidic repetitive					
21	471	8.3	640	2	A54502	S antigen precursor					
22	452	8.0	2910	2	T28156	DNA-directed RNA p					
23	442.5	7.8	1999	2	AB2018	hypothetical prote					
24	435	7.7	1192	2	A71623	probable secreted					
25	430.5	7.6	4550	2	T18440	hypothetical prote					
26	391.5	6.9	1217	2	S52714	sericin1B - silkw					
27	388.5	6.9	695	2	S27390	calcium-binding pr					
28	381.5	6.8	955	2	T18435	hypothetical prote					
29	379.5	6.7	3724	2	T18427	hypothetical prote					

30	371.5	6.6	1305	2 T00670	probable inositol
31	370	6.6	1063	2 D86731	hypothetical prote
32	368	6.5	1664	2 T18262	S-layer protein -
33	366	6.5	3844	2 T18402	asparagine/asparta
34	364.5	6.5	792	2 T42963	hypothetical prote
35	363.5	6.4	1360	2 T18403	asparagine/asparta
36	362	6.4	1193	2 G71605	hypothetical prote
37	351.5	6.2	1428	2 T08852	lustrin A - Califo
38	345.5	6.1	2481	2 D90011	FmcB protein (lmpo
39	343.5	6.1	1658	2 S55101	hypothetical prote
40	338	6.0	4776	2 E95206	cell wall surface
41	336	6.0	2485	1 H71621	serine/threonine-s
42	335.5	5.9	4910	2 S64942	probable membrane
43	333.5	5.9	1072	2 A86827	hypothetical prote
44	333	5.9	797	2 A36811	hypothetical prote
45	331.5	5.9	1419	2 T30531	agglutinin-like ad

ALIGNMENTS

RESULT 1

T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:98261511; PMID:9596732
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <N1>
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 100.0%; Score 5646; DB 2; Length 1092;
Best Local Similarity 100.0%; Pred. No. 3.3e-211;
Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MINKKNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGHNEAKAEENSVDVK	60
DB	1	MINKKNLLTKKPIANKSNKYAIRFTVGTASIVIGATLLFGLGHNEAKAEENSVDVK	60
QY	61	DSNTDDELSDNDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKESDRTS	120
DB	61	DSNTDDELSDNDQSSDEEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKESDRTS	120
QY	121	TTNVDENEATFLQKTPQDNTHLTREEVKSSSVSSSSIDTAQPSHTTINREESVQTS	180
DB	121	TTNVDENEATFLQKTPQDNTHLTREEVKSSSVSSSSIDTAQPSHTTINREESVQTS	180
QY	181	DNVEDSHVSDPFANSKIKESNTESGKEENTIEQKNVKESTTSQPSGYTNIDKISNQDE	240
DB	181	DNVEDSHVSDPFANSKIKESNTESGKEENTIEQKNVKESTTSQPSGYTNIDKISNQDE	240
QY	241	LLNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSIITEGYDDSE	300
DB	241	LLNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQGSNNVHLIKVTDQSIITEGYDDSE	300
QY	301	GVTKAHDAAENLIYDVTFEVDVKSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGIIIA	360
DB	301	GVTKAHDAAENLIYDVTFEVDVKSGDTMTVDIDKNTVPSDLTDSFTPIPKIKNSGIIIA	360
QY	361	TGYDNNKQIITFTDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKLTALSSVVK	420
DB	361	TGYDNNKQIITFTDYDVKYENIKAHKLTSYIDKSKVPNNNTKLDVEYKLTALSSVVK	420
QY	421	TTITVEYQRPNEKNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGSGT	480
DB	421	TTITVEYQRPNEKNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGSGT	480

```
QY 481 IIDDSTIIKVKVGNQNLPSNRRIYDYSEVEDVTNDYAOAGNNNDVNFNGNIDSPYI 540
Db 481 IIDDSTIIKVKVGNQNLPSNRRIYDYSEVEDVTNDYAOAGNNNDVNFNGNIDSPYI 540
QY 541 IKVLSKIDPNKDDYTTIOQTVMOTTINEYTGERTASYDNTIAFSSSGOGGDLPEK 600
Db 541 IKVLSKIDPNKDDYTTIOQTVMOTTINEYTGERTASYDNTIAFSSSGOGGDLPEK 600
QY 601 TYKIGDYVWEDVDKDGIOQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEDKYQFDGLKNG 660
Db 601 TYKIGDYVWEDVDKDGIOQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEDKYQFDGLKNG 660
QY 661 LTYKIFTEPGYPTTLKHSNTNPALESGNSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Db 661 LTYKIFTEPGYPTTLKHSNTNPALESGNSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
QY 721 VNYDTNKDGIQDDDEKGI SGVKVTLKDENGNIISTTTTDENGKYQFONLANSNVIHFDK 780
Db 721 VNYDTNKDGIQDDDEKGI SGVKVTLKDENGNIISTTTTDENGKYQFONLANSNVIHFDK 780
QY 781 PSMGTQTTTDSGDDDEQADGEEVHVITTDHDDFSIDNGYDDDESDDSDSDSDSDSD 840
Db 781 PSMGTQTTTDSGDDDEQADGEEVHVITTDHDDFSIDNGYDDDESDDSDSDSDSDSD 840
QY 841 SDSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 841 SDSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
QY 901 SDSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Db 901 SDSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
QY 961 SDSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
Db 961 SDSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
QY 1021 SDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Db 1021 SDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
QY 1081 LGKRRKNRKNKN 1092
Db 1081 LGKRRKNRKNKN 1092

RESULT 2
T28680
fibrinogen-binding protein homolog - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C;Accession: T28680
R;Johansson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A;Title: Three new members of the serine-asparilate repeat protein multigene family of Staphylococcus aureus
A;Reference number: 220510; MUID:99098700; PMID:9884231
A;Accession: T28680
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1166 <JOS>
A;Cross-references: EMBL:AJ005647; NID:el1318793; PID:el1318794; PIDN:CAA06652.1
A;Genetics:
A;Gene: sdrE

Query Match 47.8%; Score 2698.5; DB 2; Length 1166;
Best local similarity 48.2%; Pred. No. 3.3e-97;
Matches 595; Conservative 158; Mismatches 267; Indels 215; Gaps 26;

QY 1 MINKN-NLTKKPIANKSNKAIKFTVGTASIVIGATLLPGLGHNEAKAENSVDV 59
Db 1 MINKN-NLTKKPIANKSNKAIKFTVGTASIVIGATLLPGLGHNEAKAENSVDV 59
QY 60 KDSNTDDELSDNQSDDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKRSDRTE 119
Db 60 KDSNTDDELSDNQSDDEKNDVNNQSIINTDDNNQIIKKEETNNYDGIKRSDRTE 119
```

```
Db 59 ENAKQDDATTSDNKEVSVETENNSTTNNSTNP-----IKKE--TNTDSQPEAKKESTS 110
QY 120 STTNVDENEATFLQKTPQDNTHLTBEVKESS---SSVESNSSIDTAQOPSHHTINREES 176
Db 111 SSTQKQNNVATATETKPN--LEKENVXPSTDKTATEDTSVILEKKAAPNNT--ANDVT 166
QY 177 VQTSNDNVEDSHVDFANSKIKESNT---ESGKEENTIEOPNKVKEDSTTSQPSGYTNIDE 233
Db 167 TKFS-----TSEPSTSEIQTKPTTPQESTNIENSQOPQPTPSKVD---NQVTDATNPKE 216
QY 234 KIS-NODELLNLP-----INVENKARPLSTTSQAOPSIKRV-----TVNQLAAEQ 277
Db 217 PNVVSKEELKNPEKELKELVRNDSNTDHSKPVATAPTSVAPKRVNAKRFVAQPAAVA 276
QY 278 GSNVNLHILKVTDOSITIEGYDDSGVKAHAENLIYDVTFEVDDKVKSGDGTMTVDIDKNT 337
Db 277 SNNVNDLILKVTKQIKVG--DGKNVAAAHGDKDIEYDTEFTIDNKVKKGDTMTINVDKNV 335
QY 338 VPSDLTDSFTPIKIKONGSEIATGYDNKNKQIYTTFTDYVDKYENIKAHKLKLTSYIDK 397
Db 336 IPSDLTDKNPDIDITDPSGEVIAKGFDFKATKQITVFTDYVDKYEDIKSRLTLYSIDK 395
QY 398 SKYPANNTKLDVEYKTALSSVANKTIIVFYQRPENRNTANLQSMFTNIDTKNHTVETIYI 457
Db 396 KTVV--NETSLNLTFAATAGKETSQNTVTDYQDPMVHGDSNIQSIFTKLDKEDKQITIEQIYV 454
QY 458 NPLRYSAKETNVNISNG-----DEGSTIIDSTIIKVKVGNQNLPSNRRIYDY 508
Db 455 NPLKASATWTKVDIAGSQVDYGNIKLNGSIIIDQNTIEIKVKYKNSDQDLQPSNRRIYDF 514
QY 509 SEYEDVTND--DYAQLGNNNDVNFNGNIDSPYIIVKISKYDPNKDDYTTIOQTVMOTTI 567
Db 515 SQYEDVTSQDNKKSFSNNVATLDFGDIINSAVIIKVSQYTPSTDGELDIAQGTSWRTT-- 573
QY 568 NEYTGEPRTASVDNTIAFSTSSGOGGDL--PEPKYKIGDYVWEDVDKGIQNTNDNEKP 626
Db 574 DKY--GYNTAGYSNFTVTSNDTGGGGGTGVPKEKYKIGDYVWEDVDKGVQGTDSKEP 632
QY 627 LSNVLVTLTYPDGTSKSVRTDSDGKYQFGLKNGLTYYKITFTETPEGYTPTLKHSGTNPAL 686
Db 633 MANVLVTLTYPDGTSKSVRTDANGHYEFGGLKDGETTYTVKFTGYLFTKNGITDGEK 692
QY 687 DSEGNVWVTINGQDDMTIDSGFYQTPKYSIGNYVWVTNKGIOGDDSKGISGVKVTULK 746
Db 693 DSGSSVTVKINGKDDMSLDGTGYKEPKYKNGLYVWEDTNDKGIQDANEPGIKDKVKTULK 752
QY 747 DENGNIIISTTT----- 757
Db 753 DSTGKVIIGITTTDASGKYKFTLDNGNYVTFETPAGYTPVKNVTADDKDSNGLTTTGV 812
QY 758 ----- 757
Db 813 IKDADNMTLDRGFYKTPKYSGLGVWYSDNKGDKQDSTEKGIKDVTVTLQNEKEVIGTT 872
QY 758 -TDENGKYQFONLANSNVIHFDPKSGMTQTTTSDSDGDDDEQADGEEVHVITTDHDDPSI 816
Db 873 KTDENGKYRFDNLDSGKYKVIPEKPAGLTQVTNTTDED--KDADGGEVDVITTDHDDFTL 931
QY 817 DNGYVWEDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 876
Db 932 DNGYFEEDT-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 980
QY 877 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 936
Db 981 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1040
QY 937 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 996
Db 1041 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1100
QY 997 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1055
Db 1101 SDSDSDAGKHTPVKPMSS-----TTKDHINKAKALPE 1131
```


[illegible]

Db 253 LPSQTQNLVNAQGNIIAKGYDSKNTTTFYTNVDQYTNVSGSPQVAFAPKRENAATD 312
Qy 404 NTKLDVYKATLSSVNTKIIVYQRPENRNTANLQSMFTNIDTKNHTVEQTIYIN-PLRY 462
Db 313 KTAIYMEVTLGNDTYSKQVLDVY---GNQKGQQLISSTNYINNEDLSRWNTVYVNPQKKT 369
Qy 463 SAKETNV-NISGNGDEGSTIIDSTIIKVKYVGNQNLPLDSNRIRYDYSEYEDVTND-DYA 520
Db 370 YTKETFTVNLN-----GKYNPDAKFNKIEVTDQNGQVDS-FTPDTSKLDKVTGQGFVFI 423
Qy 521 QLGNNNDVINPGN-----IDSPYIIKVIKVDNPKDDYTIIOQTVTMTQTTINNYTCEFT 576
Db 424 YSNDKNTATVDLLNGQSSSKQYIIQVAYPDNSSTNGKIDYLTETQNGKSSWSN---- 479
Qy 577 ASYDNTTIAFTSSGQGGDLPEPKYKIGYVWEDVDKGIQNTNDNEKPLSNVLTLTY 636
Db 480 -SYSNVNGSSTANGD-----QKYNLGYVWEDTNKQKQDA--NEKIGKGVYILKD 529
Qy 637 PDGTS-KSVRTDEDKYQFGLKNGLYKTIPTPEGYTPTLKHSGTNPALDSGNSVWV 695
Db 530 SNGKELDRITTTDENGKYQFTGLSNG-TYSVEFSTPAGYPTTANAGTDDAVDSGLTTTG 588
Qy 696 TINGQDDMTYDSGFYQTPKYSGLNYYWYDNTKGIQGDDEKIGISGVKVTILKDENGHIST 755
Db 589 VIKADANMTLDSGFYKTPKYSGLGYVWYDNTKGIQGDDEKIGISGVKVTILKDENGHIST 648
Qy 756 TTTDENGKYQFNLNSGNYIVHPDKSGMTQTTTDSGDDDEQDADGEVHVHTITDHDDES 815
Db 649 TETDENGKYQFNLNSGNYIVHPDKSGMTQTTTDSGDDDEQDADGEVHVHTITDHDDES 707
Qy 816 IDNGYYDDESDS 875
Db 708 LDNGYVEEET-SDS 766
Qy 876 DS 935
Db 767 DS 826
Qy 936 DS 995
Db 827 DS 886
Qy 996 DS 1055
Db 887 DSDSDSDAGKHTTPKPSWTV-----KQHKHTAKALPE 918
Qy 1056 TGANEDYSGKGLTGLTGLTGLTGLTGLTGLTGLTGLTGLTGLTGLTGLTGLTGLTGLT 1089
Db 919 TGSNNNNNSNGTLFGGLFAALGSLLLFGRKQKQNK 953
RESULT 6
T28679
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28679
R:Josefsson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of *S. aureus*
A:Reference number: 220510; MUID:99098700; PMID:9894231
A:Accession: T28679
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA06651.1
C:Genetics:
A:Gene: sdrD
Query Match 33.3%; Score 1880; DB 2; Length 1315;
Best Local Similarity 36.0%; Pred. No. 1.7e-65;
Matches 498; Conservative 177; Mismatches 346; Indels 362; Gaps 46;

[illegible]

Query Match 33.3%; Score 1880; DB 2; Length 1315;
Best Local Similarity 36.0%; Pred. No. 1.7e-65;
Matches 498; Conservative 177; Mismatches 346; Indels 362; Gaps 46;

Db 409 YIKGYQDK-----IESSGKVSATDTKLRIFEVNDTSKLSDSYYADPNDSNLKVTDFKN 464
Qy 504 RIYDYSEYEDVNDYDIAQLGNNDNINFGNIDSPYIIKIVSKYDPNKKDDVTT--IOQTV 561
Db 465 RIY-----HPNVASIKFGDITKTYVVLVEGHYDNTGKMLKTQVIGENV 509
Qy 562 TMOQTINYEGETFRTASVDNTIAFSTSGGQGGDLPEPKYKIGDYVWEDVDKGIQNTN 621
Db 510-----DPTVNRDYSI--FGW-----NN 524
Qy 622 DNEKPLSNVLVTLTYPDGTSKSVRTDEGKYQFGLKXGLTKYKITFTPEGYPTLKHSG 681
Db 525 EN-----VRYGGGSA-----DG-----DSAVN-----PKDPTF-----548
Qy 682 TNPALDSEGSVWVVTINGQDDMTIDSGFYQPKYSLGNVYVDTNKDGIQDDEKGISGV 741
Db 549 -GPPVDPESP-----DPE-----561
Qy 742 KVTLDKENGNIISTTTDENGKYQFDNLNGNIVYVHFDKPSGTMQTITTDGDDDEQADG 801
Db 562-----PEPTDPEPSPDPEPSPDPP 584
Qy 802 EEVHTITDHDPSIDNGYVDRSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861
Db 585-----DSDSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 624
Qy 862 DS 921
Db 625 DS 684
Qy 922 DS 981
Db 685 DS 744
Qy 982 DSDL 1041
Db 745 DS 804
Qy 1042 G-----NSSDKSTK-----DKLPDTGANEDYSGKGTLLGTLFAGLGAL 1079
Db 805 RVTPPNNEQKAPNPKGEVNHNSKVKSHQKTDALPETG-DKSENTNATLFGAMWALLGSL 863
Qy 1080 LL-CKRKNK 1091
Db 864 LLFRKRQDHKEK 876

RESULT 10
AG1671
Probable peptidoglycan bound protein (LPXTG motif) lin1913 [imported] - Listeria innocua
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AG1671
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-882 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97143.1; PID:g16414414; GSPDB:GN00178
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin1913

Query Match 13.9%; Score 785; DB 2; Length 882;
Best Local Similarity 35.1%; Pred. No. 2.2e-23;
Matches 151; Conservative 130; Mismatches 91; Indels 58; Gaps 9;

Qy 662 TYKITFTPEGYPTLKHSGT-----NPAIDSEGSVWVVTINGQDDMTI 705
Db 95 TYPYFELSELTSILSNPNKANTKIDYKIAVLGIGNIGLFGQG-----TVNGSS-----145
Qy 706 DSGFYQPKYSLGNVYVDTNKDGIQDDEKGISGVK-----VTLKDENGNI 752
Db 146 NNFFIDTSRNAIGARV---NHLGV-----GVSSVSTFTLTIDLLALGVLTALPSANDGK 196
Qy 753 ISITTTDENGKYQFDNLNGNIVYVHFDKPSGTMQTITTDGDDDEQADGEEVHVITDH 811
Db 197 LDFARTGGLGLVDLLNSN-----AARGFITTDVGADADADADADADADA 247
Qy 812 DPFSDINGYVDRSDS 871
Db 248 DA 307
Qy 872 DS 931
Db 308 DA 367
Qy 932 DS 991
Db 368 DA 427
Qy 992 DS 1051
Db 428 DA 486
Qy 1052 KLPTGANED 1061
Db 487 ADADADADAD 496

RESULT 11
AG1299
Probable peptidoglycan bound protein (LPXTG motif) lmo1799 [imported] - Listeria monocytogenes
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AG1299
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1299
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-903 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99877.1; PID:g16411253; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1799

Query Match 13.9%; Score 782.5; DB 2; Length 903;
Best Local Similarity 35.4%; Pred. No. 2.8e-23;
Matches 154; Conservative 131; Mismatches 83; Indels 67; Gaps 10;

Qy 662 TYKITFTPEGYPTLKHSGTNPALDSEGSVWVVTINGQDDMTIDSGFYQPKYSLGNV 721
Db 94 TYPYFELP-----SELSTLSNPINRANKID---YKIAVLGGIGTG 133
Qy 722 WYDTNKGIGQ-----DDEKGISGVKVT---LKDENGNIISTTTTDE-----760
Db 134 LF--NQGTVNGSSSNFFIDSSRNAIGAKVNHLLGVGVGSVSTFTLTIDLLALGVLTALPS 191
Qy 761 --NGKYQF-----DNLNSGNIYVHFDKPSGTMQTITTDGDDDEQADGEEVHV 806
Db 192 ANDGKLDFAARTGDLGLDVLNLSN-----AARGFITTDVGADADADADADADA 242

Db	3377	ENFDNNNEGTL 3387	:
RESULT 15			
T17451			
fimbriae-associated protein Fap1 - Streptococcus parasanguinis			
C;Species: Streptococcus parasanguinis			
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000			
C;Accession: T17451			
R;Wu, H.; Fives-Taylor, P.			
submitted to the EMBL Data Library, October 1998			
A;Description: Nucleotide sequences of the fap1 locus.			
A;Reference number: Z18788			
A;Accession: T17451			
A;Status: preliminary; translated from GB/EMBL/DDBJ			
A;Molecule type: DNA			
A;Residues: 1-2570 <WUH>			
A;Cross-references: EMBL:AF100426; NID:g3929311; PID:g3929312; PIDN:AAC79868.1			
A;Experimental source: strain FW213			
C;Genetics:			
A;Gene: fap1			
Query Match 9.7%; Score 549; DB 2; Length 2570;			
Best Local Similarity 24.5%; Pred.No. 1.1e-13;			
Matches 254; Conservative 209; Mismatches 384; Indels 190; Gaps 32;			
Qy	71	SNQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKRESDRTESTTNVDENEAT 130	:
Db	63	SGNVFADE---TVLAKETLITTTDANEV-KLSSENF-----SEKAEKISLSQSES- 110	:
Qy	131	FLOKTPQDNTHLTREEVKSSSVESSENSIDTAQQPSHTTINREESVQTSNDVEDSHVSD 190	:
Db	111	-----ASESVSESIS-ESVSESSTSESSES-SESV--SESISES-VSE 151	:
Qy	191	FANSKIKESNTEGKEENTIEQPNKVKEDSTTS-----OPSGYTNIDEKISNODELLNL 244	:
Db	152	SISESISESVSESTSTSVLSESGAAGNATKGTTEKQDSVRENLDKMSI-EAEVLN- 209	:
Qy	245	PINEYENKARPLSTTSQAQPSIKRVTYNQLAAEQGSNNHLIK--VTDQSIITEGYDDSEGV 302	:
Db	210	-----DMAARKLITLDAEQOLE-----LMKSLVATQSOLEA----- 240	:
Qy	303	IKAHDAENLIYDVFVEDDKVKSGDWTVDIDKNTPSDLTDSFTTPKIKDN-SGEIIA- 360	:
Db	241	-----TKNLIGDPNATVAD-----LQIAYTTLGN--TQALGNELIKLNPNGQIYAV 285	:
Qy	361	-----TGTYDNKNKQITVFTDYVDKYENIKAHLKLTYSI-----DKSKV 400	:
Db	286	LNNTEASRAATLRSSTTGT-----KTTFTISDFSGG-----TQYWAGGNANL 330	:
Qy	401	PNNNTKLDVEYKTALSVNKTITVEYQRPNEENTAMLSMFT---NIDTKNHTVEQTIIY 456	:
Db	331	KNPISISAVYDSATGKI--SWIVEVDPTTILKSPALKTKLTGTGIYIDTSSDSKLSPT 388	:
Qy	457	INPLRYSAKETNNVINGNDEGSTIIDDSIIKIVYK-----VGDQNQLPD----- 501	:
Db	389	NVLIDGAATNPVTNPFYNGSGKIEYVSGTGTGKHTITFDTAFGSRANDLADLEIKML 448	:
Qy	502	-SNRIYDYSEYEDVTDNDYLAQLGNNDVNFNGNIDSPYIIKVIKVDPNKODYTTIQOT 560	:
Db	449	AATTLSDPHFYEKSGKNGYR-----YNGQTAPYVIA-----NDSGTAI--- 487	:
Qy	561	VTMQTTINEYTGFRFTASVNTITAFSTSGQGQDLPEKTYK-IGDYVWEDVDKDGION 619	:
Db	488	-----GGYQVSGVNADSIPTDTSQSASAKSESTKSISESIVIESSESIVGS 536	:
Qy	620	TNDNEKPLSNVLVTLTPDGTSKSVRTDEDKYQFDGLKNGLTYKIFTETPEGYTPTLKH 679	:
Db	537	VSESVSESSESSESITESSESVS-----ESISESVSESSESSESSESSESVSSE 587	:
Qy	680	SGTNPALDSGNSWWVTINGQDDHTSDSGFYQTPKYSLGNVWVVDNKGDIQGDDEKGIS 739	:

Search completed: October 5, 2004, 20:53:24
Job time : 56.7921 secs

1040 DLGNSDDKSTKDKLPDT 1056
945 -VSESVSESISESVSES 960

588 SISESVSESISESVSESISESVSESISESVSESISESVSESISESVSE 647
740 GVKVTLKDENGNIISTTTTDENGKYQFQNLNNGNIVIHFDKPSGMTQTTTDSGDDDEQDA 799
648 SVSESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISES 704
800 DGEEVHVTTITDHDHDFSIDNGYYDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 859
705 VSESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISESVSES 764
860 DS 919
765 VSESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISES 824
920 DS 979
825 ISESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISES 884
980 DS 1039
885 ISESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISESVSESISES 944

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen, Ltd.
OM protein - protein search, using sw model
Run on: October 5, 2004, 20:49:18 ; Search time 128.509 Seconds
(without alignments)
2734.476 Million cell updates/sec
Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKKNLLTKKKPIANKSN.....FAGLGALLGKKRKNKN 1092
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1351062 seqs, 321799191 residues
Total number of hits satisfying chosen parameters: 1351062
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	15	US-10-378-674-8
2	4485	79.4	930	12	US-10-615-383-10
3	4485	79.4	930	16	US-10-689-082-10
4	4485	79.4	930	16	US-10-690-184-10
5	4309	76.3	892	16	US-10-661-809-21
6	4307	76.3	892	12	US-10-282-1222A-70481
7	3351.5	59.4	670	12	US-10-282-1222A-70444
8	2787	49.4	549	15	US-10-378-674-9
9	2624.5	46.5	560	15	US-10-378-674-2
10	2586	45.8	1141	12	US-10-282-1222A-70251
11	2339.5	41.4	1633	12	US-10-282-1222A-70437
12	2303.5	40.8	1742	12	US-10-615-383-4
13	2303.5	40.8	1742	16	US-10-689-082-4
14	2303.5	40.8	1742	16	US-10-690-184-4
15	2234	39.6	1385	12	US-10-282-1222A-44324

16	2205.5	39.1	1920	12	US-10-282-122A-71413	Sequence 71413, A
17	2093.5	37.1	953	12	US-10-282-122A-44457	Sequence 44457, A
18	2069	36.6	1349	9	US-09-815-242-5898	Sequence 5898, Ap
19	2069	36.6	1349	9	US-09-815-242-13137	Sequence 13137, A
20	2011	35.6	932	9	US-09-815-242-5578	Sequence 5578, Ap
21	2011	35.6	932	9	US-09-815-242-12438	Sequence 12438, A
22	1672	29.6	343	15	US-10-378-674-4	Sequence 4, Appl1
23	1634	28.9	1021	9	US-09-815-242-5471	Sequence 5471, Ap
24	1634	28.9	1021	9	US-09-815-242-12544	Sequence 12544, A
25	1629.5	28.9	936	8	US-08-781-986A-5249	Sequence 5249, Ap
26	1629.5	28.9	936	12	US-10-329-624-5249	Sequence 5249, Ap
27	1581	28.0	935	12	US-10-282-122A-44326	Sequence 44326, A
28	1562	27.7	316	15	US-10-378-674-6	Sequence 6, Appl1
29	1536	27.2	877	12	US-10-282-122A-70428	Sequence 70428, A
30	1441.5	25.5	841	9	US-09-815-242-5779	Sequence 5779, Ap
31	1441.5	25.5	841	9	US-09-815-242-12751	Sequence 12751, A
32	1121.5	19.9	513	12	US-10-282-122A-59154	Sequence 59154, A
33	1023	18.1	265	13	US-10-073-256-78	Sequence 78, Appl1
34	865	15.3	1113	9	US-09-815-242-5836	Sequence 5836, Ap
35	836	14.8	997	9	US-09-815-242-12769	Sequence 12769, A
36	782.5	13.9	903	12	US-10-282-122A-60847	Sequence 60847, A
37	708.5	12.5	1253	14	US-10-363-798-2	Sequence 2, Appl1
38	671.5	11.9	487	12	US-10-615-383-14	Sequence 14, Appl1
39	671.5	11.9	487	16	US-10-689-082-14	Sequence 14, Appl1
40	671.5	11.9	487	16	US-10-690-184-14	Sequence 14, Appl1
41	624.5	11.1	2271	12	US-10-282-122A-43924	Sequence 43924, A
42	615.5	10.9	2283	14	US-10-172-502-4	Sequence 4, Appl1
43	607	10.8	2344	9	US-09-815-242-12713	Sequence 12713, A
44	594	10.5	1831	12	US-10-282-122A-71033	Sequence 71033, A
45	554	9.8	406	16	US-10-451-467A-262	Sequence 262, App

ALIGNMENTS

RESULT 1

US-10-378-674-8
; Sequence 8, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PARTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGATI
; FILE OF INVENTION: STAPHYLOCOCCAL PROTEINS
; FILE REFERENCE: P07556US01/BAS
; CURRENT FILING DATE: 2003-03-05
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; TYPE: PRT
; LENGTH: 1092
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-8

Query Match	100.0%	Score	5646	DB	15	Length	1092
Best Local Similarity	100.0%	Pred. No.	3.8e-307				
Matches 1092	Conservative	0	Mismatches	0	Indels	0	Gaps
0							
Qy	1	MINKKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVK	60				
Db	1	MINKKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVK	60				
Qy	61	DSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESDRTES	120				
Db	61	DSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQIIKKEETNNYDGIKESDRTES	120				
Qy	121	TTNVDENEATFLQKTPQDNTHLTFEEVKGSSSVSSNSSIDTAAQPSHTTINREESVQTS	180				
Db	121	TTNVDENEATFLQKTPQDNTHLTFEEVKGSSSVSSNSSIDTAAQPSHTTINREESVQTS	180				
Qy	181	DNVEDSHVDSFANSKIKESNTESGKEENTIEQPNKVEDSTTSQPSGYTNIDKISNQDE	240				

```
Db 181 DNVEDSHVDFANSKIKESNTSGKEENTIEQPNKVKEDSTTSQSPGYTNIDKISNQDE 240
Qy 241 LNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQSNVNHILKVTQDSITGEYDDSE 300
Db 241 LNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQSNVNHILKVTQDSITGEYDDSE 300
Qy 301 GVIKAHAENLIYDVTFFVDDKVKSGDTWTVDDIDKNTVPSDLTDSFTPIPKIKDNGSBEIIA 360
Db 301 GVIKAHAENLIYDVTFFVDDKVKSGDTWTVDDIDKNTVPSDLTDSFTPIPKIKDNGSBEIIA 360
Qy 361 TGTIDNKNKQIYTTFTDVKYENIKAHKLTSYIDSKVPNNNTKLDVEYKTLSSVNK 420
Db 361 TGTIDNKNKQIYTTFTDVKYENIKAHKLTSYIDSKVPNNNTKLDVEYKTLSSVNK 420
Qy 421 TITVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLYSKAKETNVNISGNGDSGST 480
Db 421 TITVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLYSKAKETNVNISGNGDSGST 480
Qy 481 IIDSTIIKVKVGNQNLPSNRRIYDYSEYEDVTNDYDQALGNNDVNIINFGNIDSPYI 540
Db 481 IIDSTIIKVKVGNQNLPSNRRIYDYSEYEDVTNDYDQALGNNDVNIINFGNIDSPYI 540
Qy 541 IKVISKYPDKDDYTTIOQTVTMQTTINEYTGERTASYDNTIAFSTSSGCGQDLPEK 600
Db 541 IKVISKYPDKDDYTTIOQTVTMQTTINEYTGERTASYDNTIAFSTSSGCGQDLPEK 600
Qy 601 TYKIGDYVWEDVDKQDQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDDEGKYQFDGLKNG 660
Db 601 TYKIGDYVWEDVDKQDQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDDEGKYQFDGLKNG 660
Qy 661 LTYKIFTEPGYPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Db 661 LTYKIFTEPGYPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Qy 721 VYDITNKGIOGDDKEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDK 780
Db 721 VYDITNKGIOGDDKEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDK 780
Qy 781 PSGMTQTITDSDGDDDEQADGEEVHVITIDHDDPSIDNGYVDDSDSDSDSDSDSDSDSD 840
Db 781 PSGMTQTITDSDGDDDEQADGEEVHVITIDHDDPSIDNGYVDDSDSDSDSDSDSDSDSD 840
Qy 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Qy 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Db 901 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
Qy 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
Db 961 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1020
Qy 1021 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Db 1021 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1080
Qy 1081 LGKRRKRNKN 1092
Db 1081 LGKRRKRNKN 1092
```

RESULT 2

```
US-10-615-383-10
; Sequence 10, Application US/10615383
; Publication No. US20040038327A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS EPIDERMIDIS
; FILE REFERENCE: P06335US03/BAS
; CURRENT APPLICATION NUMBER: US/10/615,383
```

```
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-615-383-10
```

```
Query Match 79.4%; Score 4485; DB 12; Length 930;
Best Local Similarity 81.1%; Pred. No. 2,4e-247;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;
```

```
Qy 4 KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYKDSN 63
Db 2 KKNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDYKDSN 61
Qy 64 TDDELSDSDQSDDEKNDVINNNQSDINTDDNNQIIKKEETNNYDGIKRSSEDRSTTN 123
Db 62 MDDELSDSDQSDNEEKNDVINNSQSINTDDNQ-IKKEETNSDAIENRSDITQSTTN 120
Qy 124 VDNEATFLQKTPQDNTLTEREVKSSVESNSIDTAQOPSHTTINREBSVQTSNV 183
Db 121 VDNEATFLQKTPQDNTLQKKEVPESSVSSMDTAQOPSHTTINSEASIQTSNE 180
Qy 184 EDHVSDFANSKIKESNTSGKEENTIEQPNKVKEDSTTSQSPGYTNIDKISNOBELN 243
Db 181 ENSRVDFANSKIIESTESNKEENTIEQPNKVRDSTTSQSPSYKNIIDEKISNOBELN 240
Qy 244 LPINEVENKARPLSTTSQAQPSIKRVTVNQLAAEQSNVNHILKVTQDSITGEYDDSEGV 303
Db 241 LPINEVENKARPLSTTSQAQPSIKRVTVNQLAAEQSNVNHILKVTQDSITGEYDDSEGV 300
Qy 304 KAHDAENLIYDVTFFVDDKVKSGDTWTVDDIDKNTVPSDLTDSFTPIKIDNSGEIATGT 363
Db 301 KAHDAENLIYDVTFFVDDKVKSGDTWTVDDIDKNTVPSDLTDSFAIPKIDNSGEIATGT 360
Qy 364 YDNKNKQIYTTFTDVKYENIKAHKLTSYIDSKVPNNNTKLDVEYKTLSSVKNKTI 423
Db 361 YDNKNKQIYTTFTDVKYENIKAHKLTSYIDSKVPNNNTKLDVEYKTLSSVKNKTI 420
Qy 424 VEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLYSKAKETNVNISGNGDSGST 483
Db 421 VEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPLYSKAKETNVNISGNGDSGST 480
Qy 484 DSTIIKVKVGNQNLPSNRRIYDYSEYEDVTNDYDQALGNNDVNIINFGNIDSPYI 543
Db 481 DSTIIKVKVGNQNLPSNRRIYDYSEYEDVTNDYDQALGNNDVNIINFGNIDSPYI 540
Qy 544 ISKIDPNKDDYTTIOQTVTMQTTINEYTGERTASYDNTIAFSTSSGCGQDLPEKTYK 603
Db 541 ISKIDPNKDDYTTIOQTVTMQTTINEYTGERTASYDNTIAFSTSSGCGQDLPEKTYK 600
Qy 604 IGDYVWEDVDKQDQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDDEGKYQFDGLKNG 663
Db 601 IGDYVWEDVDKQDQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDDEGKYQFDGLKNG 660
Qy 664 KITFETPEGYPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTPKYSIGNY 723
Db 661 KITFETPEGYPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTPKYSIGNY 720
Qy 724 DTNKDGIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSG 783
Db 721 DTNKDGIQGDDEKIGSVKVTLDKENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSG 780
Qy 784 WQQTITDSDGDDDEQADGEEVHVITIDHDDPSIDNGYVDDSDSDSDSDSDSDSDSDSD 843
```



```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-690-184-10

Query Match      79.4%; Score 4485; DB 16; Length 930;
Best Local Similarity 81.1%; Pred. No. 2.4e-242;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

QY 4 KKNLLTKKPIANKSKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVQKDSN 63
DB 2 KKNLLTKKPIANKSKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVQKDSN 61
QY 64 TDELSLSDQSSDEEKNDVNNQSINTDDNNOI IKKEETNNYDGLKESEDETESTTN 123
DB 62 MDELSLSDQSSDEEKNDVNNQSINTDDNNOI IKKEETNSNDALENRSKOITQSTTN 120
QY 124 VDENEATFLOKTPQDNTLHTEEEYKSSSVSSNSSIDTAQPSHTTINREESVQTSNV 183
DB 121 VDENEATFLOKTPQDNTLHTEEEYKSSSVSSNSSMDTAQPSHTTINSEASITSDNE 180
QY 184 EDHSVDFANSKIKESNTEGKEENTIEQPNKVEDSTTSQPSGYTNIDKISNOBELLN 243
DB 181 ENSRVDFANSKIKESNTEGKEENTIEQPNKVEDSTTSQPSGYTNIDKISNOBELLN 240
QY 244 LPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSNNVHLIKVTDQSITTEGYDDSEGI 303
DB 241 LPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSNNVHLIKVTDQSITTEGYDDSEGI 300
QY 304 KAHAENLIIYDVTPEVDKVKSGDVTMVIDKNTVPSDLTDSFTPIKIDNSGEIATGT 363
DB 301 KAHAENLIIYDVTPEVDKVKSGDVTMVIDKNTVPSDLTDSFTPIKIDNSGEIATGT 360
QY 364 YDNKAKIITFTDYDVKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALSSVKNKTI 423
DB 361 YDNKAKIITFTDYDVKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALSSVKNKTI 420
QY 424 VEYQPNENRTANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGEGSTIID 483
DB 421 VEYQPNENRTANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNGEGSTIID 480
QY 484 DSTIIKYKVGDNQNLPSNRIDYSEYEDVTNDYLAQLGNNDVNFNIDSPYIIKV 543
DB 481 DSTIIKYKVGDNQNLPSNRIDYSEYEDVTNDYLAQLGNNDVNFNIDSPYIIKV 540
QY 544 ISKYDPNKDDYTTIQQTVTMTQTTINEYTGFRFASVDNTIAFTSSGQGGDLPPKTYK 603
DB 541 ISKYDPNKDDYTTIQQTVTMTQTTINEYTGFRFASVDNTIAFTSSGQGGDLPPKTYK 600
QY 604 IGDYVWEDVDKGIQNTNDNEKPLSNVLTLPDGTGSKSVRTDEGKYQFDGLKNGLT 663
DB 601 IGDYVWEDVDKGIQNTNDNEKPLSNVLTLPDGTGSKSVRTDEGKYQFDGLKNGLT 660
QY 664 KITFETPEGYPTPLKHSNTPALDSEGNVWVTINGQDDMTIDSGFYQTPKYSILGNVWY 723
DB 661 KITFETPEGYPTPLKHSNTPALDSEGNVWVTINGQDDMTIDSGFYQTPKYSILGNVWY 720
QY 724 DTNKGIOGDEKGI SGVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIYHFKPSPG 783
DB 721 DTNKGIOGDEKGI SGVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIYHFKPSPG 780
QY 784 MTQTITDSGDDDEQADAGEVHVHTITDHDHDFSIDNGYDDSDSDSDSDSDSDSDSDS 843
DB 781 MTQTITDSGDDDEQADAGEVHVHTITDHDHDFSIDNGYDDSDSDSDSDSDSDSDS 833
QY 844 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 903
DB 834 ----- 833
QY 904 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 963
```

```

DB 834 ----- 833
QY 964 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1023
DB 834 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 880
QY 1024 GSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1083
DB 881 -----NSSDKNTKDKLPDTGANEDHDSKGTLLGALFAGLGALLGK 921
QY 1084 RKRRKRNKN 1092
DB 922 RKRRKRNKN 930

RESULT 5
US-10-661-809-21
; Sequence 21, Application US/10661809
; Publication No. US20040101919A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; FILE REFERENCE: P07741US01/BAS
; CURRENT APPLICATION NUMBER: US/10/661,809
; PRIOR FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/410303
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-661-809-21

Query Match      76.3%; Score 4309; DB 16; Length 892;
Best Local Similarity 80.5%; Pred. No. 1.5e-232;
Matches 847; Conservative 19; Mismatches 26; Indels 160; Gaps 3;

QY 41 LFLGLGHNEAKAENSVDVQKSDTDDLSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 100
DB 1 LFLGLGHNEAKAENSVDVQKSDTDDLSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 59
QY 101 KEETNNYDGLKESEDETESTTNVDENEATFLOKTPQDNTLHTEEEYKSSSVSSNSSI 160
DB 60 KEETNSNDALENRSKOITQSTTNVDENEATFLOKTPQDNTLHTEEEYKSSSVSSNSSI 119
QY 161 DTAQPSHTTINREESVQTSNVDSHVSDFPANSKIKESNTEGKEENTIEQPNKVEDS 220
DB 120 DTAQPSHTTINSEASITQSDNEENSRSVDFANSKIKESNTEGKEENTIEQPNKVEDS 179
QY 221 TTSQPSGYTNIDKISNOBELNLPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSN 280
DB 180 ITSQPSGYTNIDKISNOBELNLPINEYENKARPLSTTSAQPSIKRVTYNQLAAEQGSN 239
QY 281 VNHLIKVTDQSTTEGYDDSEGIKAHAENLIIYDVTPEVDKVKSGDVTMVIDKNTVPS 340
DB 240 VNHLIKVTDQSTTEGYDDSDGIIKAHAENLIIYDVTPEVDKVKSGDVTMVIDKNTVPS 299
QY 341 DLTDSTFTPIKIDNSGEIATQTYDNKNKQIYVTFDVTYDVKYENIKAKHLKLTYSIDKSKV 400
DB 300 DLTDSTFAIPKIDNSGEIATQTYDNKNKQIYVTFDVTYDVKYENIKAKHLKLTYSIDKSKV 359
QY 401 PNNNTKLDVEYKTALSSVKNKTIITVEYQPNENRTANLOSMTNIDTKNHTVEQTIYINPL 460
DB 360 PNNNTKLDVEYKTALSSVKNKTIITVEYQPNENRTANLOSMTNIDTKNHTVEQTIYINPL 419
QY 461 RYSAKETNNVNSGNGEGSTIIDSTIIKYKVGDNQNLPSNRIDYSEYEDVTNDYLA 520
DB 420 RYSAKETNNVNSGNGEGSTIIDSTIIKYKVGDNQNLPSNRIDYSEYEDVTNDYLA 479
QY 521 QLGNNNDVNFNIDSPYIIKISKYDPNKDDYTTIQQTVTMTQTTINEYTGFRFASVD 580
```



```
Db 480 OLGNNDVNFNGNIDSPYIIKISKYDPNKKDYTTIQQVTVMTQTTINEYTGFRASYD 539
Qy 581 NTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLTLYPDGT 640
Db 540 NTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLTLYPDGT 599
Qy 641 SKSVRTDEBKQYQFDGLKNGLTYYKITPETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 700
Db 600 SKSVRTDEBKQYQFDGLKNGLTYYKITPETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 659
Qy 701 DMTIDSGFYQTPKYSIGNVVYDNTKDKGIGQDDEKIGSVKVTLDKNGNIISTTTTDE 760
Db 660 DMTIDSGFYQTPKYSIGNVVYDNTKDKGIGQDDEKIGSVKVTLDKNGNIISTTTTDE 719
Qy 761 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADAGEEHHVHTITDHDDFSIDNGY 820
Db 720 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADAGEEHHVHTITDHDDFSIDNGY 779
Qy 821 YDESQSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 880
Db 780 YDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 795
Qy 881 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 940
Db 796 ----- 795
Qy 941 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1000
Db 796 -----DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 832
Qy 1001 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060
Db 833 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 860
Qy 1061 DYCCKTLLTFLGALGKLLGKRKRKNKN 1092
Db 861 DHDSKGTLLGALFAGLALGKRKRKNKN 892

RESULT 6
US-10-282-122A-70481
; Sequence 70481, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Habelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70481
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-10-282-122A-70481

Query Match 76.3%; Score 4307; DB 12; Length 892;
Best Local Similarity 80.4%; Pred. No. 2e-232;
Matches 846; Conservative 20; Mismatches 26; Indels 160; Gaps 3;

Qy 41 LFLGLGHNEAKAENSVDYKDSNDTDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 100
Db 1 MFLGLGHNEAKAENTVDYKDSNMDDDELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 59

Qy 101 KEETNYDIEKRSERDRTSTNVNDENEATFLOKTPQDNTLHTEEFVKESSSVESNSSI 160
Db 60 KEETNSDAIENRSDITQSTTNNVDENEATFLOKTPQDNTLHTEEFVKESSSVESNSSM 119

Qy 161 DTAQOQSHHTIINREESVQTSNDVEDSHVDSFANSKIKESNTESGKEENTIEQNKVKEDS 220
Db 120 DTAQOQSHHTIINSEASIQTSNEENRSVDSFANSKILIESNTENKENTIEQNKVREDS 179

Qy 221 TTSQPSGYTNIDKISNQDELNLPINEYENKARPLSTTSAQPSIKRVTNQLAAEQSN 280
Db 180 ITSQPSYKNIDKISNQDELNLPINEYENKARPLSTTSAQPSKRVTVNQLAAEQSN 239

Qy 281 VNHLIKVTQOSTEGYDDSEGVKAHDAENLIYDVTFEVDDKVGSGDTMTVDIDKNTVPS 340
Db 240 VNHLIKVTQOSTEGYDDSDGIIKAHDAENLIYDVTFEVDDKVGSGDTMTVDIDKNTVPS 299

Qy 341 DLTDSFTPIKIDNSGEIATGTYDNKNKQITVFTDYVDYKYNKIAHLKLTYSYIDKSKV 400
Db 300 DLTDSFAIPKIDNSGEIATGTYDNKNKQITVFTDYVDYKYNKIAHLKLTYSYIDKSKV 359

Qy 401 PNNNTKLDVEYKTALSSVANKTITVEYQRPNEANTANLQSMFTNIDTKNHTVEQTYIINPL 460
Db 360 PNNNTKLDVEYKTALSSVANKTITVEYQRPNEANTANLQSMFTNIDTKNHTVEQTYIINPL 419

Qy 461 RYSAKETNVNISNGDEGSTIIDDSITIKYKVGDNQNLPSDSNRIYDYSEYEDVTNDYYA 520
Db 420 RYSAKETNVNISNGDEGSTIIDDSITIKYKVGDNQNLPSDSNRIYDYSEYEDVTNDYYA 479

Qy 521 QLGNNDVNFNGNIDSPYIIKISKYDPNKKDYTTIQQVTVMTQTTINEYTGFRASYD 580
Db 480 QLGNNDVNFNGNIDSPYIIKISKYDPNKKDYTTIQQVTVMTQTTINEYTGFRASYD 539

Qy 581 NTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLTLYPDGT 640
Db 540 NTIAFSTSSGQGGDLPPPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLTLYPDGT 599

Qy 641 SKSVRTDEBKQYQFDGLKNGLTYYKITPETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 700
Db 600 SKSVRTDEBKQYQFDGLKNGLTYYKITPETPEGYTPTLKHSGTNPALDSEGNVWVINGQ 659

Qy 701 DMTIDSGFYQTPKYSIGNVVYDNTKDKGIGQDDEKIGSVKVTLDKNGNIISTTTTDE 760
Db 660 DMTIDSGFYQTPKYSIGNVVYDNTKDKGIGQDDEKIGSVKVTLDKNGNIISTTTTDE 719

Qy 761 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADAGEEHHVHTITDHDDFSIDNGY 820
Db 720 NGKYQFDNLNSGNYIVHFDKPSGWTQTTTDSGDDDEQADAGEEHHVHTITDHDDFSIDNGY 779
```


Best Local Similarity 98.78; Pred. No. 9.6e-148;
Matches 542; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy	52	EENSVDVQKSDNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQI1KKEETNNYDIE	111
Db	1	EENSVDVQKSDNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQI1KKEETNNYDIE	60
Qy	112	KRSERDTESTTNDVENEATFLOKTPQDNTLTHTEEVKESVSSSSSITDAQPSHTTI	171
Db	61	KSEDRTESTTNDVENEATFLOKSPQDNTLTHTEEVKESVSSSSSITDAQPSHTTI	120
Qy	172	NREESVQTSNDVSDSHVSPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQSGVTNI	231
Db	121	NREESVQTSNDVSDSHVSPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQSGVTNI	180
Qy	232	DEKISNODELNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQS	291
Db	181	DEKISNODELNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLIKVTDQS	240
Qy	292	ITEGYDSDSEGVKAHAENLIYDVTFEVDDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKI	351
Db	241	ITEGYDSDSEGVKAHAENLIYDVTFEVDDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKI	300
Qy	352	KDNGSEIATGTVDNKNKQIITYTFTDVVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEY	411
Db	301	KDNGSEIATGTVDNKNKQIITYTFTDVVDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEY	360
Qy	412	KTALSSVKNKTIITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLYSKAKETNNVI	471
Db	361	KTALSSVKNKTIITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLYSKAKETNNVI	420
Qy	472	SGNGDEGSTIIDSTIIKVKYKVDNQLPDSNRIDYSEYEDVTNDDYAOQGNNDVNIIN	531
Db	421	SGNGDEGSTIIDSTIIKVKYKVDNQLPDSNRIDYSEYEDVTNDDYAOQGNNDVNIIN	480
Qy	532	FGNIDSPYIIKISKYDPNKDDYTTIOQTVMQTINEYTGEPRTASVDNTIAFSTSSGQ	591
Db	481	FGNIDSPYIIKISKYDPNKDDYTTIOQTVMQTINEYTGEPRTASVDNTIAFSTSSGQ	540
Qy	592	QGQDLPEK 600	
Db	541	QGQDLPEK 549	

RESULT 9

US-10-378-674-2
; Sequence 2, Application US/10378674
; Publication No. US20040006209A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; TITLE OF INVENTION: MONOCLONAL AND POLYCLONAL ANTIBODIES RECOGNIZING COAGULASE-NEGAT
; FILE REFERENCE: P07556US01/BAS
; CURRENT APPLICATION NUMBER: US/10/378,674
; CURRENT FILING DATE: 2003-03-05
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-378-674-2

Query Match 46.58; Score 2624.5; DB 15; Length 560;
Best Local Similarity 92.34; Pred. No. 1.2e-138;
Matches 512; Conservative 17; Mismatches 25; Indels 1; Gaps 1;

Qy	46	HNEAKAEENSVDVKDSNTDDELSDNSQSSDEEKNDVINNNQSIINTDNNQI1KKEETN	105
Db	7	HHHGGSENTVDVKDSNNDDDELSDNSQSSDEEKNDVINNNQSIINTDNNQI1KKEETN	65

Qy	106	NYDGIKRSERDTESTTNDVENEATFLOKTPQDNTLTHTEEVKESVSSSSSITDAQOQ	165
Db	66	SNDAIENRSDITQSTTNDVENEATFLOKTPQDNTLTHTEEVKESVSSSSSMDTAQOQ	125
Qy	166	PSHTTINREESVQTSNDVSDSHVSPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQ	225
Db	126	PSHTTINSEASIQTSNDSEENSRVSDPFANSKILBSNTESNKEENTIEQPNKVEDSITSQ	185
Qy	226	SGYTNIDEXISNODELNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLI	285
Db	186	SSYKNIDEXISNODELNLPINEYENKARPLSTTSQAQPSIKRVTVNQLAAEQGSNNVHLI	245
Qy	286	KVTDQSIITEGYDSDSEGVKAHAENLIYDVTFEVDDKVKSGDTMTVDIDKNTVPSDLTDS	345
Db	246	KVTDQSIITEGYDSDSDGIIKAHAENLIYDVTFEVDDKVKSGDTMTVDIDKNTVPSDLTDS	305
Qy	346	FTTPKIKDNGSEIATGTVDNKNKQIITYTFTDVVDKYENIKAHKLKTSYIDKSKVPNNNT	405
Db	306	FAIPKIKDNGSEIATGTVDNTNKQIITYTFTDVVDKYENIKAHKLKTSYIDKSKVPNNNT	365
Qy	406	KLDVEYKTALSSVKNKTIITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLYSK	465
Db	366	KLDVEYKTALSSVKNKTIITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLYSK	425
Qy	466	ETNVNISGNGDEGSTIIDSTIIKVKYKVDNQLPDSNRIDYSEYEDVTNDDYAOQGN	525
Db	426	ETNVNISGNGDEGSTIIDSTIIKVKYKVDNQLPDSNRIDYSEYEDVTNDDYAOQGN	485
Qy	526	NDVINFGNIDSPYIIKISKYDPNKDDYTTIOQTVMQTINEYTGEPRTASVDNTIAF	585
Db	486	NDVINFGNIDSPYIIKISKYDPNKDDYTTIOQTVMQTINEYTGEPRTASVDNTIAF	545
Qy	586	STSSGQGGDLPEK 600	
Db	546	STSSGQGGDLPEK 560	

RESULT 10

US-10-282-122A-70251
; Sequence 70251, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70251
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-70251

Query Match      45.8%; Score 2586; DB 12; Length 1141;
Best Local Similarity 46.7%; Pred. No. 4e-136;
Matches 577; Conservative 151; Mismatches 267; Indels 240; Gaps 27;

QY 1 MINKGN-LLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDV 59
Db 1 MINRDNKKAITKGMISNRLKFSIRKYVTGASILVGTTLIFGLGNQEAKAENT--ST 58

QY 60 KDSNTDDELSDNSQSDDEEKNVDVNNOSINTDDNNQIIKKEETNNVDGIEKESDRTE 119
Db 59 ENAKQDATTSDNKEVSETEENSTTENDSNP-----IKKE--TWDSQPEAKEESTT 110

QY 120 STTNVDENEAFTLOKTPDONTNHLTEEBVKES--SSVESNSSSIDTAQPSHTTINREES 176
Db 111 SSTQQQNNVTATTETKPN--IEKENVKPSTDKTATEDTSVILEEKKAPNYT--NNDVT 166

QY 177 VQSDNVEDSHVDFANSKIKESNT---ESKENTIEOPNKVKEDSTTSQPSGYTNIDE 233
Db 167 TKPS-----TSEIQTKPTTPOESTNIENSQPTPSKVD--NQVTDATNPKE 211

QY 234 KIS-NODELLMLP-----INEYENKARPLST--TSAQP-----SIKVTVNOLAAEQ 277
Db 212 PVNVSKELKNPEKELVENDNNTDRSYKPVATATSVAPKELNAKMPFAVAQPAVA 271

QY 278 GSNVHLIKVTDQISITGYDDSEGVIIKAHAENLIYVTFEVDKVKSGDTMTVDIDKNT 337
Db 272 SNNVNDLITVTKTIKVG-DGKDNVAAAHGDKOIEYDTEFTIDNKKVKGDTMTINVDKNV 330

QY 338 VPSDLTDSFTIPKIKNSGETIANGTYDNKNKQITVFTDYVDKYNENIKAHKLTSYIDK 397
Db 331 IPSDLTQNDPIDTDSGEVIAKTKATKQITVFTDYVDKIEDIKARLTLSYIDK 390

QY 398 SKVNNNTKLDVEYKTALSSNKNITITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYI 457
Db 391 QAVP-NETSLNLTATAGKETSQNVSDYQDPWVHGDSNIQSIFTKLDENKQIEQIYV 449

QY 458 NPLRYSAKETNVNISNG-----DEGSTIIDSTIIKVKVKGDNQMLPDSNRIYDY 508
Db 450 NPLKKTATNTKVDITAGSQVDYDGNIKLNGSTIIDQNTKVKVKNPNOQLPQSNRIYDF 509

QY 509 SEYEDVTND--DYAQLGNNDVNIINFGNIDSPYIIKIVSKYDPNKNDDYTTIQQVVTMTTI 567
Db 510 SQVEDVTSQDNKKSFPNNVATDFGINSAYIIKVVSKTPTPSDGEIDTAQGTSMRTI-- 568

QY 568 NEYTGERTASYNNTIAFTSSGQGDLL--PPEKTYKIGDYVWEDVDKGIQNTNDNEKP 626
Db 569 DKY--GYNYAGYSNFIVTSNDTGGDGTVPKEKLYKIGDYVWEDVDKGVQGTDSKEP 627

QY 627 LSNVLVTLTYPDGTSKSVRTDEGKYQFDGLKNGLYTKITFETPEGYPTPLKHSIGNPAL 686
Db 628 MANVLVTLTYPDGTSKSVRTDANGHYEFGLGDKDGETYTVKFPETPAGYLPKTVNGTTDGEK 687

QY 687 DSEGSVWVVTINGQDD----- 702
Db 688 DSGSSITVKGKDDMSLDTGFKPKYNLGDVWEDNKNKGIQANBPGIKDKVKVTK 747

QY 703 ----- 702
Db 748 DSGTKVIGTTTTDASGKYKFTDLNNGNYTFEFPAGYTPTVKNTTAAEDKSDNGLTTTGV 807
```

```
703 -----MTIDSGFYOTPKYSIGNYVWYDTNKGIGQDDEKIGSGVKTLLKDENGNIISTTT 756
808 IKDADNWTLDGFKYKPKYSLGYYVWYDNGKQKQSTKGIKDVKTLLNEKEVIGTT 867
757 TTDENGKYQPDNLNSGNYIVHFQKPSGTMQTTTDSGDDDEQDADGEEVHTVITDHDDFS 816
868 KTDENGKYRFDNLDSGKYKVIPEKPAGLTQVNTTTEDD--KDADGGEVDVTTDHDFFIL 926
817 DNGYDDDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 876
927 DNGYFEEDT-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 955
877 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 936
956 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1015
937 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 996
1016 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1075
997 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1055
1076 SDSDSDAGKHTPVKPS-----TTKDHNNKAKALPE 1106
1056 TGANEDYGSKGTLLGLTFAGLGAALL--GKRRKNRK 1089
1107 TGSENNNGNATLFGGLFAALGSLLLFGRKKQNK 1141

RESULT 11
US-10-282-122A-70437
; Sequence 70437, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```


US-10-615-383-4

Query Match 40.8%; Score 2303.5; DB 12; Length 1742;
Best Local Similarity 34.6%; Pred. No. 4.2e-120;
Matches 605; Conservative 146; Mismatches 313; Indels 685; Gaps 46;

Qy 2 INKQNLTKKPTANKSKNAIRKFTVGTASIVIGATLLFGLGHNAKAEEN-----S 55
Db 18 INKRVDFL-----SNKVKYSIRKFTVGTASILVATLMPGAADNEAKAEDNQLESAS 71

Qy 56 VQDVKDS--NTDDELS--DSNDQSDDEKNDVNN-----NQSN 91
Db 72 KEOGGRDNENSKLQVLDLNGSHSEKTTNNVATEVKKVEAPTTSVSKPKANEAVV 131

Qy 92 TOD-----NNQII---KBEETNNYDGEK-----RSEDRTSTNVN 127
Db 132 TNESTKPTTEAPTVEESIAETPKTTTQODSTKKNPSPILKONLNSSTTSKESKTD 191

Qy 128 EATFLQ-KTPQDNTHLTEBEVK-BSSSVESNSSIDTAQOPS-----167
Db 192 STKQAMSTNKSNIJDTNDSPTQSEKTSQANNDSTNQSAFQKSPSEQVKYTKFN 251

Qy 168 -----HTTIN--REESVOTSDNVEDSHVSDPFANSKIKESNTESGKEENTIPQPNKVK 218
Db 252 DEPTQDVEHTTKLTSVSTSSVNDK--QDYTRSAV-----ASLGVDNSMETAITNAVR 305

Qy 219 DSTTSQSPSGYTNIDEKI-----SNQDELLNLP--INEYENKARPLSTTSQAQSIKR 267
Db 306 DNLDLKAASREQINEAIIAEALKKDFNSPDYDVTPLALNRSQSKNSP--HKSASP--R 360

Qy 268 VTNVQLAAE--QGSNVHLIKVTDQ--SITEGYDDSEGVKAHDAENLIYDVTFEVDKV 323
Db 361 MNLMSLAEEPNSGKRVNDKVTNPTLSLKNSSNHNANNVWPTSEQFNLKANYELDDSI 420

Qy 324 KSGTMTVDIDKNTVPSLTDSTFPIPKIKONSGEIIATGYDNKNKQITTYFTDYVDKY 393
Db 421 KEGDTFIKGYIRPGGLEPAIKTQURSKDGSIVANGVDKTTNTTYFTFYVDQYQ 480

Qy 384 NIKAKHLKLTSYDKSKVPNNNTKLDVEYKLTALSSVKNKTIIVEXORPNNRNTANLQSNFTN 443
Db 481 NITGSFLLIATPKRETAIKONQNPMEVTIANEVVKDFIVDYGNKKDNITT--AAVAN 537

Qy 444 IDTKNHTVEQTIINPLR-----YSAKETN-----468
Db 538 VDNVNNKHENVYILNQNQNPKYAKYFTVKGGEFIPGEVKVYEVTDNAMVDSFNPDLN 597

Qy 469 -----VNISGNGD-----EG 478
Db 598 SSVNVDVTSQFAPKVSADGTRVDINFARSMAWGGKVIYVQAVRPTGTGVNVTYWLTRDG 657

Qy 479 STIIDD-----STII-----KVYKVGDD-----NQN-----498
Db 658 TTNNDPVRGKTSTTVYLNKSGSTAQDNFTYSLGDYVWLDKKNKGVDQDDDEKGLAGVYV 717

Qy 499 LPDSN-----RI-----YDYSEYEDVT-----NDYIAQLGNNDVN-----529
Db 718 TLKDSNNRELQRTVTDQSGHYQFDNLQNGTYTVEFAIPDNYTSPANNSTNDAIDSDGER 777

Qy 530 -----INFGNIDSPYI--IKVTSKY-----DPNKO-----DYVTIQ 558
Db 778 DGRKVVVAKGTIN--NADNMTVDGTFLTPKXNVGVWEDTNKGIQDNDKEGSGVK 835

Qy 559 QTV-----TWQTTINBYTGEFTASVDN--TIAFST-----SSQGGQ--DLPE 599
Db 836 VTLKNKNGDTIGITTTDSNGKYEFTGLENGDYITIEFETPEGYTPTKQNSGSDGKDSNGT 895

Qy 600 K-----TYKIGDYVWEDVDKGIQNTNNEKPLSNVJLTLYPDG 639
Db 896 KTTVTVKADANKTIDSGFYKPTNGLGVWEDTNKGIQ--DDSEKGISGVKVTLDKNG 953

Qy 640 TS-KSVRTDGDGYQFDGLKNGLYTKITFTPEGYTPTLKHSGTNPALDSEGNVWVTIN 698
Db 954 NAIGTITTDASGHYQFKLENG-SYTVFETPSPGYPTKANSQDITVDSNGITTTGIIN 1012

RESULT 13

US-10-689-082-4

; Sequence 4, Application US/10689082

; Publication No. US20040142348A1

; GENERAL INFORMATION:

; APPLICANT: FOSTER, Timothy

; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS

; FILE REFERENCE: P06335US04/BAS

; CURRENT APPLICATION NUMBER: US/10/689,082

; CURRENT FILING DATE: 2003-10-21

; PRIOR APPLICATION NUMBER: 09/386,962

; PRIOR FILING DATE: 1999-08-31

; PRIOR APPLICATION NUMBER: 60/098,443

; PRIOR FILING DATE: 1998-08-31

; PRIOR APPLICATION NUMBER: 60/117,119

; PRIOR FILING DATE: 1999-01-25

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4


```
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-689-082-4

Query Match      40.8%; Score 2303.5; DB 16; Length 1742;
Best Local Similarity 34.8%; Pred. No. 4.2e-120;
Matches 605; Conservative 146; Mismatches 313; Indels 685; Gaps 46;

Qy 2 INKNNLLTKKPKPIANKSKYAIRKETVTGASIVIGATLLFGLGHNEAKAEEN-----S 55
Db 18 INKRVDFL-----SNKNKYKIRKPTVGTASILVGLTLMFGAADNEAKAEADNQLESAS 71

Qy 56 VQDVKDS--NTDDELS--DSNDQSSDEEKVDVNN-----NQSNIN 91
Db 72 KBEQKGRDNENSKLQVLDLNGSHSEKTTNNVNAFEVKVREAPTTSVDSKPKANEAVV 131

Qy 92 TDD-----NQI1---KEETNNYDGEK-----RSEDRTSTTNVDEN 127
Db 132 TNESTKPKTEAPTVEESIAETPKTSTTQODSTERNKPNLKNLSSSTTSKESKTDEH 191

Qy 128 EATFLQ-KTPQDNLHTEBEVK-ESSSVESNSSIDTAQOPS-----167
Db 192 STKQAMSTNKSNDLTDNDSPQSEKTSQANNDSTNQSAKQDLSKPEQKVYKTKFN 251

Qy 168 -----HTTIN-REESVOTSDNVESHVSDFANSKIKESNTESGKBENTIEOPKNVKE 218
Db 252 DEPTQDVEHTTTLKTPSPVSTGSDVNDK--QDVTSAV-----ASLGVDSNETEAITNAVR 305

Qy 219 DSTTSQPSGYTNIDEKT-----SNQDELNLP--INEYENKARPLSTTSAQPSIKR 267
Db 306 DNLDLKAASREQINEAIIAEALKKQFSDPDYGVDTPLALNRSQKNSP--HKSASP---R 360

Qy 268 VTNQLAAE--QGSNNVHLKVTDO--SITEGVDDSEGVKAHDAENLIYDVTFEVDVKV 323
Db 361 MNLMSLAAEPNSGNVNDVKIINPTLSLKNHNNHANNVWPTSNEQFNLKANYELDDSI 420

Qy 324 KSGDWTVDIDKNVPSDLTDSFTPIPKIKDNGSEIATGYDNKNKQITVTFDYVDKYE 383
Db 421 KEGDPTIKYQVIRPGGLEPAIKTOLRSKGSIVANGVYDKTNTVTFYTNVYDQYQ 480

Qy 384 NIKAHKLTSYIDKSKVPNNKTKLDVEYKTSALSVANKTIIVEYQRPNEARTANLQSNFTN 443
Db 481 NITGSFDLIATPKRETAIKDNQNPMEVTTIANEVVKQDFIVDYGKKNKNTT---AAVN 537

Qy 444 IDTKNHTVEQTIYINPLR-----YSAXETN-----468
Db 538 VDNVNNKHNEVVLNQNQNPYKAKYFSTVNGSEFIPGEVKYVEVTTNMAVDSFNPDLN 597

Qy 469 -----STI1-----KVKVGD-----VNISNGD-----EG 478
Db 598 SSNVKDVTSQFAPKVSADGTRVDINFAFARMANGKKYIVTQAVRPTGTGNVYTYEYWLTRDG 657

Qy 479 STI1DD-----STII-----KVKVGD-----NON-----498
Db 658 TTNNDPFRGKSTTVTYLNGSSTAQGDNPNTYSLGDYVWLDRKNGVQDDDEKGLAGVYV 717

Qy 499 -LPDSN-----RI-----YDYSEYEDVT-----NDYQAOLGNNDVN-----529
Db 718 TLKDSNNRELQVTTDSQHYQFNQNLQNGYTYVEFAIPDNTYTPSPANNSTNDALDSGCR 777

Qy 530 -----INFGNIDSPYI---IKVISKY-----DPNKO-----DYTTIQ 558
Db 778 DGRKVVVAKGTIN--NADNMVTDGTFLTPKYNVGVVWEDTKDGIQDDNEKSGISGVK 835

Qy 559 QTV-----TWQTTINYEGBFRTASVDN---TIAFST-----SSCGQG-DLPPE 599
Db 836 VTLANKNGDITGTTTDSNGKYEFTGLENGDYTIEFETPEGTPTPKQNSGSDGKDSNGT 895

Qy 600 K-----TYKIGVYVWEDVDKQIGINTNDNEKPLSNVLTLYTPDG 639
Db 896 KTTVTVKADANKTIDSGFYKPTYNLIGVYVWEDTNKQIGQ--DDSEKGISGVKVTLDKNG 953
```

RESULT 14

```
US-10-690-184-4
; Sequence 4, Application US/10690184
; Publication No. US20040141997A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: METHODS FOR TREATING OR PREVENTING INFECTIONS FROM COAGULASE-
; FILE REFERENCE: P06335US05/BAS
; CURRENT APPLICATION NUMBER: US/10/690,184
; PRIOR FILING DATE: 2003-10-21
; PRIOR APPLICATION NUMBER: 09/386,962
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 5, 2004, 20:44:57 ; Search time 31.9642 Seconds
(without alignments)
1763.712 Million cell updates/sec

Title: US-10-806-288-15
Perfect score: 5646
Sequence: 1 MINKNNLLTKKKPIANKSN.....FAGLGALLGKKRKNKNKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4490	79.5	930	4	US-09-134-001C-5314
2	4485	79.4	930	4	US-09-386-962C-10
3	2698.5	47.8	1166	4	US-09-200-6508-7
4	2303.5	40.8	1742	4	US-09-386-962C-4
5	1982	35.1	930	4	US-09-200-6508-3
6	1880	33.3	1315	4	US-09-200-6508-5
7	1629.5	28.9	936	4	US-08-956-171B-5249
8	1628	28.8	933	3	US-08-233-728-2
9	1628	28.8	933	3	US-09-421-868-2
10	1613.5	28.6	918	4	US-09-200-6508-1
11	671.5	11.9	487	4	US-09-386-962C-14
12	635.5	11.3	2137	4	US-09-134-001C-4463
13	488	8.6	1027	4	US-08-956-171B-5254
14	482	8.5	466	4	US-09-134-001C-4749
15	408.5	7.2	669	4	US-09-107-532A-6532
16	377.5	6.7	1183	4	US-09-134-001C-3530
17	349	6.2	345	3	US-08-856-253-7
18	331.5	5.9	1161	4	US-09-327-536-2
19	315.5	5.6	1112	2	US-08-714-402-2
20	301.5	5.3	1231	3	US-08-904-263A-4
21	301.5	5.3	1231	4	US-09-434-123A-4
22	295	5.2	2315	4	US-09-543-681A-5434
23	292.5	5.2	886	4	US-08-956-171B-5235
24	292.5	5.2	2504	4	US-09-328-352-5821
25	286	5.1	1060	4	US-08-911-393-2
26	282	5.0	3052	2	US-08-557-122A-26
27	282	5.0	3052	4	US-09-262-666-26

28	279	4.9	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
29	278	4.9	699	4	US-09-134-001C-4054	Sequence 4054, Ap
30	276.5	4.9	1085	1	US-08-431-080-28	Sequence 28, Appl
31	276.5	4.9	1085	2	US-08-938-534-28	Sequence 28, Appl
32	276.5	4.9	1085	4	US-09-345-294-28	Sequence 2, Appl
33	276	4.9	2391	2	US-08-446-855A-2	Sequence 2, Appl
34	276	4.9	2391	3	US-09-150-741-2	Sequence 2, Appl
35	271	4.8	1115	2	US-08-568-459A-2	Sequence 2, Appl
36	271	4.8	1115	2	US-08-487-826B-2	Sequence 2, Appl
37	271	4.8	1115	4	US-09-210-288-2	Sequence 2, Appl
38	271	4.8	1115	6	5198347-6	Patent No. 5198347
39	270.5	4.8	1833	4	US-08-621-944A-4	Sequence 4, Appl
40	270.5	4.8	1833	4	US-08-945-567D-4	Sequence 4, Appl
41	270.5	4.8	1992	4	US-08-621-944A-3	Sequence 3, Appl
42	270.5	4.8	1992	4	US-08-945-567D-3	Sequence 3, Appl
43	269.5	4.8	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
44	269	4.8	60	4	US-09-386-962C-19	Sequence 19, Appl
45	267	4.7	2123	3	US-08-968-685A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match	79.5%	Score 4490;	DB 4;	Length 930;
Best Local Similarity	81.2%	Pred. No. 1.1e-241;		
Matches	884;	Conservative 19;	Mismatches 26;	Indels 160; Gaps 3;
QY	4	KGNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVKDSN	63	
Db	2	KGNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAENSVDVKDSN	61	
QY	64	TDELSDSNDQSSDEKNDVINNNQSIINTDDNNQIIKKEETNNYDGIKRSBORTSTTN	123	
Db	62	MDELSDSNDQSSNEEKNDVINNNQSIINTDDNNQ-IKKEETNNDAIENRKSQITQSTTN	120	
QY	124	VDENEATFLQKTPQDNTLHTEEEKSSSVSSNSSIDTAQPSHTTINREESVQTSQDNV	183	
Db	121	VDENEATFLQKTPQDNTLHTEEEKSSSVSSNSSMDTAQPSHTTINSEASIQTSQDNE	180	
QY	184	EDSHVSDPFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYNIIDEKINQDELLN	243	
Db	181	ENSRVSDPFANSKIIESNTESNKEENTIEQPNKVKEDSTTSQPSGYNIIDEKINQDELLN	240	
QY	244	LPINEYENKARPLUSTTSAOPSIKRVTVNOLABQGSNNVHLIKVTQDSITEGVDDSEGI	303	
Db	241	LPINEYENKRVPLUSTTSAOPSSKRVTVNOLABQGSNNVHLIKVTQDSITEGVDDSDGI	300	
QY	304	KAHDAENLIYDVTFFEDVDKVKSGDTMTVDIKNTVPSDLTDSFTIPKIKNSGEIATGT	363	
Db	301	KAHDAENLIYDVTFFEDVDKVKSGDTMTVNDIKNTVPSDLTDSFTIPKIKNSGEIATGT	360	

Query Match	40.8%;	Score 2303.5;	DB 4;	Length 1742;
Best Local Similarity	34.6%;	Pred. No. 6.9e-120;		
Matches 605;	Conservative 146;	Mismatches 313;	Indels	685; Gaps 46
Qy	2	INKKNLLTKKKPIANKSNKVAIRKFTVGTASIVIGATILLFGLGHNEAKAEN-----S	55	
Db	18	INKRVDFL-----SNKVKYSIRKFTVGTASILVGATLMFGAADNEAKAAEDNQLSAS	71	
Qy	56	VQDVKDS--NTDDELS--DSNDQSSDEKNDVNN-----NQSN	91	
Db	72	KEEQKGRSDNENSKLNOVDLNGSHSSEKTTNNVNAATEVKVKEAPTTSVDSKPKANEAV	131	
Qy	92	TDD-----NNQI---KKEETNNVDGIEK-----RSDRTESTTNVDEN	127	
Db	132	TNESTPKTTEAPTNVNEESIAETPKTSQTQODSTEKNPNLSKONLNSSSTTSKESKTDEH	191	
Qy	128	EATFLQ-KTPQDNTHLTBEEVK-BSSSVESNSSIDTAAQPS-----	167	
Db	192	STKQAQMTNKSNDLTDNSPTQSEKTSQOANDSTDNQSASFQKLDSPSQVKYKTKFN	251	
Qy	168	-----HTTN-REESVQTSNDVEDSHVSDFPANSKIKESNTSGKEENTIEQPNKYKE	218	
Db	252	DEPTQDVHEHTTKLTPSVSTDSVNDK--QDYTRSASV-----ASLGVDNSNETAITNAVR	305	
Qy	219	DSTTSQSPGYTNIDEKI-----SNODELLANLP--INEVENKARPLSTTSQAQPSIKR	267	
Db	306	DNLDLKAASRQINEAIAEALKKDFSNPDYGVDTPLALNQSQKNSP--HKSASP--R	360	
Qy	268	VTVNQLAAE--QGSNVANHLIKVTDQ--SITBEGYDDSEGVIKAHDAENLIYDVTPEVDDKV	323	
Db	361	MNLSMLAAEPNSGKNVNDKVKITNPTLSLNSNNHANNVWPTSNQPNLKANYELDDSI	420	
Qy	324	KSGDTMTVDIDKNVPSDLTDSFIPKIKNSGIIATGTVDNKNKQITTYTFTVDYKYE	383	
Db	421	KEGDTFIKGYQYTRPGLELPAIKQLRSKDGSIANGVYDKTNTTYTFTFNVDYQYQ	480	
Qy	384	NIKAHLKLTSYIDKSQVFNNTKLDVEYKTSALSSVNKTIITVEYQKPNENRRTANLQSMFTN	443	
Db	481	NITGSPDLIATPKRETAIKDNQNPMEVTIANEVVKKDFIVDYGKKNKNTT--AAVAN	537	
Qy	444	IDTKNHTVEQTIYINPLR-----YSAKETN-----	468	
Db	538	VDNVNKHNEVYLNQNNQNPKYAKYFSTVKNGEFIFGEVKVYEVTDTNAMVDGSPNPLN	597	
Qy	469	-----VNTSGNGD-----EG	478	
Db	598	SSNVKDVTSQFAPKVSADGTRVDINFARSMANGKXYIVTQAVRPTGTVVTEWLTRDQ	657	
Qy	479	STIID--STII-----KVYKVGD-----NQ--	498	
Db	658	TTNTNDFYRGTKSTVTYLYNGSSTAQGDNPYSGLGDYVWLDKKNKGVQDDDEKGLAGVYV	717	
Qy	499	-LPDSN-----RI-----YDSEYEDVT-----NDQYAOGLGNNDVN-----	529	
Db	718	TLKDSNNRELQVTTDQSGHYQFNLQNGTVTVEFAIPDNPTSPANNSTDAIDSQER	777	
Qy	530	-----INFGNIDSPYI---IKVISKY-----DPNKD-----DYTTIQ	558	
Db	778	DGTRKVVVAAGKTIN--NADNNMTDGTFLYTPKYNVDYVWEDTNKDGIQDDNEKGI SGVK	835	
Qy	559	QTV-----TWQTTINEYTGEPRTASVDN---TIAPST-----SSGGQG-DLPPPE	599	
Db	836	VTLLKNKNGDTIGTTTDSNGKYBFTGLENGDYTIEFETPEGYTPTKQNSGDEGKDSNGT	895	
Qy	600	K-----TYKIGYVWEDVDKQIQNTNDNEKPLSNVLVTLTYTPDG	639	
Db	896	KTTVTVKDANKTIDSGFYKPTNGLDYVWEDTNKGIQ--DUSEKIGISGVKTLKDKNG	953	
Qy	640	TS-KSVRTDBDGKQFQDGLKNGLYTKITFTPTPGYTPPTLKHSGTNPALDSEGNVWVTIN	698	
Db	954	NAIGTGTITDASGHYQFQGLENG-SYTVFEFTPGGYTPTKANSQDQITVDSNGITGTGIIIN	1012	

RESULT 5

US-09-200-650E-3
; Sequence 3, Application US/09200650E

Patent No. 6680195
GENERAL INFORMATION:

APPLICANT: Patti, Joseph M.
APPLICANT: Foster, Timothy J.

APPLICANT: Hook, Magnus A.O.

; APPLICANT: Bidhinn, Deirdre N.

APPLICANT: Perkins, Samuel L.

TITLE OF INVENTION: Extracellular

FILE REFERENCE: P06283US2/BAS
CURRENT APPLICATION NUMBER: US/09/200,650E

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

; PRIOR APPLICATION NUMBER: 60/066,815

; PRIOR FILING DATE: 1997-11-26

; PRIOR APPLICATION NUMBER: 60/098,427

; PRIOR FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 23

```

; SOFTWARE: PatentIn Ver. 2.0

```



```
; SEQ ID NO 3
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-3

Query Match
 35.1%; Score 1882; DB 4; Length 930;
Best Local Similarity 42.1%; Pred. No. 2.2e-102; Indels 184; Gaps 27;
Matches 460; Conservative 156; Mismatches 293;

Qy 2 INKNNLLTKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVDQVKD 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MNKKATATNRKGMIPNRLNKFIRKYSVGTASILVGTLLFGLSGHEAKAE-----52

Qy 62 SNTDDELSDNQSSDEEKNDVINNQSINTDDNNQIIKKEETNNYDGIKRESDRTST 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 -HTNGELNQSKNETTAPSENKTKKVDLSRLKDNQ-----T 88

Qy 122 TNVDENATFLQKTPQDNTHLTREEVKESSESSVSIDTAQPSHTTINREESVQTS 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
89 ATADQPKVT-----MSDSATVKETSSNM--QSPQATANQSTTTKTSNV 129

Qy 182 NVEDSHVDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSOPSGYTNIDKISNDEL 241
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
130 TTNDKSTTSYNETDKSLNQA-----KDVSTI-----157

Qy 242 LNLPINEYENKARPLSTTSQPSIKRVTVNQLAA-EQGSNNHLIKVT--DQSITEGY-D 297
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
158 -----PKTTIKRPTLRMAVNTVAAPQOGTVNVDKVPESNIDIAIDKGVN 204

Qy 298 DSEGVK--AHDENLIYDVFEDDKVKSQDVTMVDIDKNTVPSDLTDSFTPIKIKONS 355
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
205 OTTGKTEFWATSSDVLKLNKANYTIDDSVKEGDTFTFKYGYFRGSLRPSQTNLYNAQ 264

Qy 356 GEIATGYDNKKNQIYTFYDVKYENIKAKHLKLTSDYDKSKVPNNNTKLDVEYKTA 415
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 GNIAKGIYDSTWYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTY 324

Qy 416 SSVNKTITVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYIN-PLRYSAKETNV-NISG 473
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 DTYSEBIIVDY--GNKKAQPLISSYNNINDELNRNTAYVQPKNTYTKQTFVNTLT- 380

Qy 474 NGDEGSITIDSTIIKVKYKQDNQL-----PDSNRIYDSEYEDVTNDDYQALGNNDY 528
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
381 ----GYKFNENAKNFKIYEVTDQNFVDSFTPTSKLVDTPQDFVI---YSNDKNTATV 433

Qy 529 NINFGNIDS--PYIIKVI-----SKYDPNKDDYTTIQQVTVMOTINEXTGEFTASYDN 581
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 DLMKGQTSNNKQYIIQVAVPDSNSTDNGKIDY-----TLDTDKTKYSW---SNSYSN 483

Qy 582 TIAPSTSSGQGGDLPEKTYKIGDYVVDKQIGQNTNDNEKPLSNVLVLTLYPDGTS 641
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
484 VNGSSTANGD-----QKKYNLGDYVWEDTNKQKQDA--NEKGIKGVVYVILKDSNGKE 534

Qy 642 -KSVRTDEDGKYQDGLKNGLYTKITFETEGYTPTLKHSHTPALDSEGNVWVINGQ 700
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
535 LDRITTDENGKYQFTGLSNG--TYSVEFTAGYPTTANVTGTDVAVDSGLTTTGVKDA 593

Qy 701 DDMTIDSGFGYQPKYSLGNVYVDYTNKDGIGQDDEKIGSVKVTLDKDENGMIISTTTDE 760
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
594 DNMTLDSGFKYTKPKYSLGDYVWYVDSNKGKRDSTKGIKGVKTLQNEKEVEIGTETDE 653

Qy 761 NGKYQFNLSNGYIYHFDKPSGMGTQTTSDGDDDEQADGEEVHVHTITDHPDPSIDNGY 820
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
654 NGKYRFDNLSGKYKVFEPAGLTGTGTTEDD--KADGGGEVDVTITDHPDFTLDNGY 712

Qy 821 YDDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 880
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
713 YEET-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 765

Qy 881 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 940
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
766 SNSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 825
```

```
Qy 941 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1000
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
826 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 885

Qy 1001 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1060
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
886 SDAGKHTPAKPMSTV-----KQHKHTAKALPETGSEN 917

Qy 1061 DYSGSKGTLGLTLF 1073
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
918 NNSNNGTLFGGLF 930

RESULT 6
US-09-200-650E-5
; Sequence 5, Application US/09200650E
; Patent No. 6680195
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Bidhinn, Deirdre Nl
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-5

Query Match 33.3%; Score 1880; DB 4; Length 1315;
Best Local Similarity 36.0%; Pred. No. 1.6e-96;
Matches 498; Conservative 177; Mismatches 346; Indels 362; Gaps 46;

Qy 1 MINKKN-NLLTKKPKTANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVDQV 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MLNRENKTAITRKMVSNRLNKFIRKYSVGTASILVGTLLFGLGQEAQAKAEASTNKL 60

Qy 60 KDSNTDDELSDNQSSDEEKNDVINNQSINTDDNNQIIKKEETNNYDGIKRESDRTE 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 NEATT-----SASDNQSSD--KVDMLQNLQEDNTKDNQ---KEM-----VSSQGNETS 105

Qy 120 STTNVDENATFLQKTPQDNTHLTREEVKESSESSVSIDTAQPSHTTINREESVQT 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 NGKLIKES--VOSTGKNKVESTAKSDQASPKSTNEDLTKQ-----TISNQALQ- 157

Qy 180 SNVEDSHVDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSOPSGYTNI--DEKISNQ 238
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
158 PDLQENKSVN-----VQPTNEENKVD-----AKTESIT-----LNVKSDAISKSN 198

Qy 239 DELL--NLPINEYENKARPLSTTSQAP-----SIKRVTVNQLAAEQGSNNHLIKYTDQSI 292
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
199 DETLDVNNNSNNENNADIILPKSTAPKRLNTRMRIAAPQSPSTEAKNVNDLITSNTL 258

Qy 293 TEGYDSDSEGVIKAHDAENLIYDVFEDDKVKSQDVTMVDIDKNTV-----PSDLTDSF 346
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
259 VVDADKKNKIVPAQDYLSLKSQIT--VDDKVKSGDYFTIKY--SDTVQVYGLNPEDIKN-- 313

Qy 347 TIPKIKD-NSGEIATGYTNKKNQIYTFYDVKYENIKAKHLKLTSDYDKSKVPNNNT 405
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
314 -IGDKDPNNGETIATAKHDTANNLITYTFTDVRFNQVGMGINYSIYMDADTIP--VS 370

Qy 406 KLDVEYKLTALSSVNKTIITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTI 455
```


US-09-200-650E-1

Query Match 28.6%; Score 1613.5; DB 4; Length 918;
Best Local Similarity 36.9%; Pred. No. 6.4e-82;
Matches 430; Conservative 127; Mismatches 281; Indels 327; Gaps 37;

QY 6 NLLTKKKPI---ANKSNKAIKFTVTGASIVIGATLLFGLGHNEAKAEENSVDVKDS 62
DB 1 NGVIFLKKRIDYLSNKKYKISRRFTVGTTSVIGATILFGIGNHQASEQS----- 53

QY 63 NTDELSDSNDQSDSEKNDVINNOQSINT--DDNQIIEKEETNNYDGIKESESDTES 120
DB 54 -NUTTQSSKNAASADSEKNNMIETPO-LNTANDTSDISANTNSANVDSTTKPMSTQTSN 111

QY 121 TTNVDENEATFLOKTPD---NTHLTBEEVKSESSVSSNSIDTTAQPSHTTINREESV 177
DB 112 TTT---TEPASTNETPQTAIKQATAAKQDQVTPQEGNSQVD-----NKTNDANSI 162

QY 178 QTSNDNVEDSHVDSPANSKIKESNTSEKENTIEQPNKVKEDSTTSQPSGYTHIDEKISN 237
DB 163 ATNSELKNSQTLDPQS-----SPQTSNAQGTSEKPSVRTRAVRSLAV 205

QY 238 QDELLNLPINEXENKARPLSTSAQPSIKRVTVNQLAAEQSNVNHILIKVTQDOSITEGYD 297
DB 206 AEPVNV-----AADAQGTNVN-----DKVTASNFX 230

QY 298 DSEGVIKAHDAENLIYDVTFEVDKVKSGDTMTV-----DID-----KNTVPSDLT 343
DB 231 LEKTTDPNPQSGNTFMAANFTVTDKVKSGDYFTAKLPDSLGTGNGVDVYSNNTMP----- 286

QY 344 DSTPIPKIKNSEGIATGYDNKNQIYTFDYDYDKYENIKAKHLKLTYSYIDKSKVPNN 403
DB 287 ----IADIKSTNGDWAKATYDLTKYTFYFDYVNNKENINGQFSLPLFTDRAPKPS 342

QY 404 NTKLDVEYKALSVNKTITVEYQRP---NENRTANLQSMFTNIDTKN--HVEQTIYI 457
DB 343 GT-YDANINIADEMFKNIITVSSPIAGIDKPKGANISSQILGVDTASQNTYKQTVFV 401

QY 458 NPLYSAKETNVNISNGDEGSTIIDDS-----TIKVKYVGNQNLPLDS----- 502
DB 402 NPKQVRLGNTWYIKGYDK-----IBESGKVSATDTKLRIFEVNDTSLKSDSVYADPND 457

QY 503 -----NRIYDYSEVEDVNDYIAQLGNNDVNFNGNIDSPYIIKIVISKYDPNK 551
DB 458 SNLKEVTDQPKNRIY-----YE-----HPNVASIKFPGDITKYTVYLVREGHYDNTG 502

QY 552 DDVTT--IQQTVTMTTINEYTGERTASYDNTAFSTSGQGGDLPPKTKYIGDYVW 609
DB 503 KNLKTQVIQENV-----VRYGGGSA-----DGVTRDYSI--FGW 527

QY 610 EDVDRKQIQNTNDNEKPLSNVLVLTYPDGTGSKSVRTDEGKYQFQGLKNGLYKITFET 669
DB 528 -----NNEN-----VVRYGGGSA-----DG-----DSAVN----- 547

QY 670 PEGYTPTKUSGNTWPAIDSEGNVWVTINGQDDMTIDSGFYQPKYSLGNVYVDTNKG 729
DB 548 PKQPTP-----GPPYDPEFSP-----DPEPEPTDPEFSPDPEFSPDPDSDSDS 595

QY 730 IQGDEKGIKGVKVTLDKENGNIISTTTTDDNGKYQFDNLNSGNYIVHFDPKPSGWTQTTT 789
DB 596 DSGSD----- 601

QY 790 DSGDDDDQADGERVHVHTITDHDPSIDNGYDDESDSDSDSDSDSDSDSDSDSDSDS 849
DB 602 DSGSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 653

QY 850 DS 909
DB 654 DS 713

QY 910 DS 969
DB 714 DS 773

QY 970 DS 1029
DB 774 DS 833

QY 1030 DS 1067
DB 834 DS 892

QY 1068 LGTLTFLAGLALL-CRRKRNK 1091
DB 893 LFGAMWALLGSLLLPRKQKHKEK 917

RESULT 11

US-09-386-962C-14
; Sequence 14, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-962C-14

Query Match 11.9%; Score 671.5; DB 4; Length 487;

Best Local Similarity 53.1%; Pred. No. 6e-30;
Matches 152; Conservative 33; Mismatches 76; Indels 25; Gaps 3;

QY 814 FSDNGYDDESDS 873
DB 17 FWTTSLSFQHAEGNHIDINFSKQIDRWAKNINRVNDTSRTGISMNSDNDLDT 76

QY 874 DS 933
DB 77 DIVNSDSENDYLDS 136

QY 934 DS 993
DB 137 DS 196

QY 994 DS 1046
DB 197 DSD-----SDS 240

QY 1047 KSTKDKLPPTGANEDYGSKTLTGLTFLAGLALLGKRRKRNK 1092
DB 241 QPQRPKYNTQNNNNINNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNH 284

RESULT 12

US-09-134-001C-4463
; Sequence 4463, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964

Qy	47	NEAKAENSVDVQKSDNTDDLSLSDNSQSDSEKRVNINQOSINTDNNQIILKKEETNN	106
Db	854	SESASTKQLSESASTSTSDSASEARKSESTKSTSLSESTSVSDSASVTSSESAST	913
Qy	107	YDGLIE-KRSEDRTESTTIVDNEATEFLQKTPQDNNHLTEEBKRESSVSESSSIDTAQQ	165
Db	914	STSVSGSTSTSIUSTSTSTSDSASIIKASEASTSKULSESVESTSDSASTSTSVSDSN	973
Qy	166	PSHTTINRESVQTSNDVESHVDFANSKIKESNTESGKEENTIEQPNVKED-STTSQ	224
Db	974	SASTSLKSTSTSVSDSTSTST-TSDSASTSTSESDSASTSLSESTSVSDSTSTST	1032
Qy	225	PSGYTNIDEKISNODELLNLPINEYENKARPLSTTSAPQIKRVTVNQLAAEGSNVHL	284
Db	1033	DSASMSASESSENSKS-----TSLSESTSTSLSGSTASTSDSASTSTSESDSTSTSL	1087
Qy	285	IKVTDQSI-----TEGYDDDSGVKAKH-----DAENLIYDVTFEYDDKV-----KSGDTMT	330
Db	1088	SESTSTSLSGSTASTSDSASTSTSESDSTSESTSLSESTSVSDSTASTSESASTST	1147
Qy	331	VDIDKNTVPSDLTDSFTPIPKIKNSGHIATGYDNNKKOITVTFDYVDKYENIKAHLK	390
Db	1148	SESENSASTSLGCSUS-TSISDST-----STSDSDSASTST-----	1184
Qy	391	LTSYIDKSKVPNNNTKLDVRYKTALESVNKTIIVVEYQRPENRTANLQSMFTNIDTKNHT	450
Db	1185	-----ESEDSTSTSLSESTSTSLSDSTSTSTSE-----SASTSTSESDST	1225
Qy	451	VEGTIYNPLRYSAKETNNVINSGBEGSTIIDDTIIKYKVGDNQNLPDNSRIYDYSE	510
Db	1226	SESTSL-----SESTSTSVS---DSTASTSDSASTST-SVSDSESASTST-----ISE	1269
Qy	511	YEDVTNDVYQALGNNDVNINFGNIDSPYIIKVIYSKVDNKDYYTIIQOTVTVMQITNEY	570
Db	1270	SLSTSVSDSTSTSTSDSASTSTSESD-----TSESTSLSESTSTSTSVSDS	1314
Qy	571	TGEFRTASVNTIATFSTSGQGQGLPPEKYKIGDYVWEDVDKGIQNTDNDNEKPLSNV	630
Db	1315	T-----SASTSDSASTSTSESES-----DSASTSLSGSTSTSTSLSDS	1350
Qy	631	LVTLYVPDGTGKSVRTDEQKQYQDGLKNGLYTKITFETPEGVTPLKHSCTNPALDSG	690
Db	1351	TSTST-SDSASTST-SESDSERASTSLSGSTSTSLSDSTST-----TSTSDSASTSTSV-SDS	1404
Qy	691	NSVWVTINGQDMDTIDSGFYQTPKYSLGNVYVDTNKGIDGDEKIGISGVKVLKDENG	750
Db	1405	NSASTSLSG-----SLSTSVSDSTSTSTSDSASASTST-----ESDSERASTSLSGSTSTST	1456
Qy	751	NIISTTTTIDENGKYQFDNLNSGNYIVHFDPKPGMTQTTTSDGDDDEQDADGEBHVITID	810
Db	1457	DSTSTSTSD-----SASTSTSV-SESENSTSTSTSTSESLSTSVSD	1493
Qy	811	-----HDDFSIDNGYIDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD	864
Db	1494	STSTSTSDSASTSTSVSDSDSASTSSESESVSTSDSESTSTSTSDSASTSTSVSESNSTST	1553
Qy	865	SD	924
Db	1554	SUGSTSTSVSDSTSTSDSASTSESDSDSASTSSSESVSTSVSDSTSTASTSESAST	1613

```
Query Match      8.6%; Score 488; DB 4; Length 1027;
Best Local Similarity 19.6%; Prad.No.2.4e-19;
Matches 232; Conservative 194; Mismatches 496; Indels 264; Gaps 41;
```

Qy 8 LLTKCKPIANKSH-KYAIRKFTVTGTSIVIGATLLFGLGHN-EAKAEENSVDQVKDMSNTD 65

Dd 1 ILHLKGDI I VKNRLRGIQRHKLGAASVFLGTWII VVGNGDKKEAASQPKTITVEENG-- 58

; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Atinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6532:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...669
; SEQUENCE DESCRIPTION: SEQ ID NO: 6532:
US-09-107-532A-6532

Query Match 7.2%; Score 408.5; DB 4; Length 669;
Best Local Similarity 23.4%; Pred. No. 3.7e-15;
Matches 188; Conservative 114; Mismatches 263; Indels 237; Gaps 35;
QY 307 DAENLIYDVT-----FEVDDKVKSGDWTVIDIKNTVPFSDLTDSFTPIKIKONS 355
Db 61 DPGNIYDNTGNYGLLGIAQFHFVAKNT--TVNAHTDGNIAHTEL-----DAKQNF 111
QY 356 GEIATGTYDNKQIYTFDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTA 415
Db 112 GTELHEGLL---NQEIYI-----QKTDSPANSGIPTGSTEMN---KPVV 151
QY 416 SSVNKTIVYQRP--NENRTANIQSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISG 473
Db 152 GS-NVDVALQGHPTVNGSRDLHTPDQFYKDRAGH-----YID---FQEFKALNVAA 202
QY 474 NGDEGSTIIDSTIIKVKVGDNQNLDP-SNRIYDYSEYED---VTNDDYAOIGNNNDV 528
Db 203 N--DLATI---TPAKYTTAAD---FPDMNRTIDLDLSDGTGLVNI DAELVTMTPL 253
QY 529 NINFGNIDSPYIIKISKYDNKDDYTIQOTVTMOTIYNEYGEFRTA----SYDNITIA 584
Db 254 QIINPN-DQVVVFNVI-----NSSALNVQSPIKYNDRSNHETEDFSDANISWFGNEMT 307
QY 585 FSTSSGQGGDLPEKTYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVLTYPDGTSKSV 644
Db 308 DLTISAPFQG-----TILAFNATIR-I 328
QY 645 RTDEGKYQFDGLANGLTYKITFETPEGYTPTLKHSCTNPALDS-----EGNSVWVWITNG 699
Db 329 TQNGDGTI-----IGTNVILDAAATNRWDPNEIFISDNG 361
QY 700 QDDMTIUSGGFYQTPKYSILGNVYVYDNTNKGIQGDEKGISGVKVTCLKDENGNIISTTTTD 759
Db 362 TD-----TTDTSD 369
QY 760 ENKGYQFDNLNGNYIVHFDKPSGMTOTTTDSDGDDQADAGEVHVITTDHDDFSDING 819
Db 370 -----STDTSSTSDS--SDASTSDSTDTTSTTSSTSDSTDSASS- 405
QY 820 YYDDSD 879
Db 406 ---DSDTDTTSD 460
QY 880 DSD 939
Db 461 SSSDSTDTTSTSDS--TUTSASSDSDTDTTSTSDSASSDSDTDTTSTSDTSDTSDSASS 519

QY 940 DSD 999
Db 520 DSDTSTTTSDS--TDTSSASSDSDTDTTSTTSDSDSDSTTTSDS--TDTSSASSDSDTSTTS 577
QY 1000 DSD 1055
Db 578 DS-SDSSTSSDSDTDTSTSTSDTRHASVFSFKTSDDNRTNENLTNKTSDNDNGNKSRAALPK 636
QY 1056 TGANEDYSGKGTLLGLTIFAGLG 1077
Db 637 TGSQSN--NWITLAGVILLVIG 656

Search completed: October 5, 2004, 20:54:15
Job time : 46.9642 secs